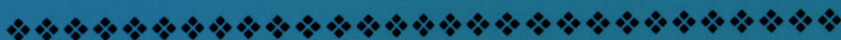


# BESZÉDKUTATÁS '99

**Elméleti és alkalmazott fonetikai  
tanulmányok**



**MTA Nyelvtudományi Intézete  
Budapest 1999**





# **BESZÉDKUTATÁS '99**

**Elméleti és alkalmazott fonetikai  
tanulmányok**

Szerkesztette:

**Gósy Mária**

A Magyar Tudományos Akadémia Nyelvtudományi Intézete  
Budapest 1999

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

Technikai szerkesztő:  
Menyhárt Krisztina

ISSN 1218-8727  
© MTA Nyelvtudományi Intézete

Felelős kiadó: Kiefer Ferenc igazgató

## Tartalom

<b>Gósy Mária – Nikléczy Péter:</b> A beszélő felismerése a beszéde alapján: elméleti háttér és módszertani megközelítések.....	1
<b>Gocsál Ákos:</b> Egyéni különbségek az artikulációs tempó percepciójában.....	19
<b>Menyhárt Krisztina:</b> Szláv anyanyelvűek magyar beszédének fonetikai sajátosságai.....	30
<b>Brenner Kolomán:</b> Magánhangzók kontrasztív akusztikai elemzése.....	44
<b>Gósy Mária:</b> A beszédprodukciónak tudatos változtatása: a beszélő személy utánzása.....	53
<b>Olaszy Gábor:</b> Beszédatadbázisok készítése gépi beszéd-előállításához.....	68
<b>Navracsics Judit:</b> „Jó lenne egy szakértőtől megtudni, hogy melyik az igazi anyanyelv.” Többyneműség és önazonosítás.....	90
<b>Lengyel Zsolt:</b> Hangzó beszéd – írott beszéd: kommunikáció.....	103
<b>Hoffmann Ildikó:</b> A mentális lexikon reprezentációja ép és Broca-afáziás személyeknél.....	113
<b>Osmanné Sági Judit:</b> A szókezdő szókiegészítő hatásának vizsgálata afáziás betegeknél .....	125



# **A BESZÉLŐ FELISMERÉSE A BESZÉDE ALAPJÁN: ELMÉLETI HÁTTÉR ÉS MÓDSZERTANI MEGKÖZELÍTÉSEK**

**Gósy Mária – Nikléczy Péter**  
**MTA Nyelvtudományi Intézete**

## **Bevezetés**

Ha hallunk egy szót, annak akusztikai hullámformája a fülön keresztül a hallóközpontba jut, majd a Wernicke-területre kerül, ahol a hangsor, majd annak szemantikai tartalma feldolgozódik. Egyúttal azonban számos más döntéssorozat is történik agyunkban. Ha a szót egy számunkra jól ismert személy ejtette ki, akkor képesek vagyunk ezt a személyt azonosítani. Ez azonban nem ilyen egyszerű és bizonyos korlátozásokkal működik. Nem mindig elegendő egyetlen szó, hogy felismerjünk valakit, ugyanakkor sokszor a telefonvonal szűk frekvenciatarományja is lehetővé teszi, hogy beazonosítsuk, ki van a vonal másik végén.

A jelen tanulmány központi kérdése az, hogy vajon a beszéde alapján felismerhető-e a beszélő személy. Az előzőekben már érintettük, hogy a felelet erre egyetlen igennél vagy nemnél lényegesen összetettebb. Az elmúlt évtizedek alatt a fonetikával, illetve beszédakusztikával foglalkozó szakemberek alapos és kimerítő vizsgálatokat végeztek a beszédelemzés legtöbb területén (vö. Stevens 1998). Sikertelen akusztikai elemzéssel mintegy rekonstruálni a beszédet, sőt – bizonyos korlátokkal – beszédfelismerő rendszerek is működnek. Azzal a ténnyel azonban, hogy az emberi hang magában rejtje az egyéni jellemzőket is, csak az utóbbi 10-15 évben kezdtek behatóbban foglalkozni. A kiinduló kérdés az volt, hogy a beszéd szegmentális vagy szupraszegmentális részében keresendő-e az egyéni hangra utaló összetevő, vagy esetleg mindkettő tartalmazza azt.

A beszédet hallgató ember képes a beszélő személy kizárólag akusztikus úton történő felismerésére, még akkor is, ha a hozzá eljutó beszéd sokszorosan torzul. (Természetesen csak akkor, ha a torzulás

mértéke egy bizonyos határt nem ér el.) A torzulás történhet (i) spektrális szinten (például az alsó vagy felső összetevők intenzitásának csökkenése következtében), (ii) szupraszegmentális szinten (például a beszéd dallamában vagy tempójában történik változás), (iii) valamennyi szinten (a spektrum nagy részét elfedő zaj esetén).

A cél azoknak a feltételeknek a meghatározása, amelyek a) lehetővé teszik, b) korlátozzák és c) nem teszik lehetővé/gátolják a beszélő személyének azonosítását.

A „közvetett” kommunikáció nagymértékű fejlődése – mindenekelőtt a telefonálásra gondolva – idézte elő azt, hogy az emberi beszéd középpontba került mint az egyén egyik legfőbb önazonossága. Számos területen vált szükségessé a személy biztos azonosítása; a biztonsági rendszerekben, a beléptető rendszerekben, banki azonosító rendszerek és a kriminalisztikában. A tudománynak arra a kérdésre kell mindenekelőtt válaszolnia, hogy vajon a beszéd valóban olyan mértékben jellemző-e az egyénre, mint az ujjlenyomat. Amennyiben e kérdésre igenlő választ ad a tudomány, a következő kérdéssorozat a beszélő azonosításának feltételeit, az azonosítás módszertani megoldásait és az azonosítás biztonsági fokának meghatározását érinti. Az alkalmazott fonetika központi kérdése pedig az, hogy melyek azok a paraméterek, amelyek kétséget kizáróan felidéznek a beszélő személyt, azaz megtörténik a beszélő személy azonosítása. Az alkalmazott fonetika új ága, az ún. törvényszéki fonetika ('forensic phonetics'), amely önálló diszciplínaként első ízben 1995-ben jelent meg a Fonetikai Világkongresszus programjában, foglalkozik a beszélőnek a beszéde alapján történő felismerésével.

Az új analízis eljárások következtében a nem kriminalisztikai célú alkalmazásokban (például banki rendszerekben) a beszélő felismerésének problémája – még telefonon át is – megoldottnak tűnik. A hetvenes évektől indultak meg az erre irányuló kutatások és fejlesztések (pl. Doddington et al. 1976), mára többféle, megbízhatóan működő rendszer létezik a világban. Némelyikük állítólag 99%-os biztonsággal képes a beszélő személy azonosítására. A kutatók különböző algoritmusok alkalmazásával vagy különféle többszörös szűrő eljárásokkal igyekeztek meghatározni a beszélő személyazonosságát. A kidolgozott eljárásokkal sikerült – technikailag jó minőségű rögzített beszéd



esetében – 90% fölötti eredményt elérni a beszélő személy azonosságának meghatározásában, de a vizsgálathoz általában 40-50 s hosszúságú hanganyagra van szükség.

Ahhoz, hogy megértsük, miért mégis megoldatlan probléma a törvényszéki beszélőfelismerés; nézzük meg, mit jelent az egyén azonosítása a biztonsági rendszerekben. A beszélő valamilyen módon azonosítja önmagát (kóddal, névvel stb.), vagyis azonnal csökkenti a lehetséges beszélő személyek számát. A beszéd alapján történő személyfelismerésnek tehát arra kell válaszolnia, hogy valóban a feltételezett személy jelentkezett be. Egy többé-kevésbé meghatározott szöveget kell a beszélőnek bemondania (például szókapcsolatot, szókapcsolatokat vagy rövid mondatot). Általában az ún. normalizált, hosszú idejű átlagos spektrumelemzéssel, amelynek során az aktuálisan bemondott szöveg különféle jegyvektorait vetik össze a beszélőtől korábban tárolt szöveg paramétereivel. Ezt követően egy hasonlósági indexet számítanak. Az egyezést a küszöbértéktől való távolság szerint határozzák meg. Ezekben az esetekben tehát a beszélő felismerését számos tényező részben megkönnyíti, részben pedig kizárólagosan lehetővé teszi. A beszélő kooperatív, ez azt jelenti, hogy azt szeretné, hogy megtörténjen a biztos azonosítása. Létezik a beszélőtől már korábban tárolt, jó akusztikai és felvételi körülmények között rögzített beszédminta. Ismert az aktuális bejelentkezés körülménye, az összevetés tehát valóban gyorsan és jó hatásfokkal elvégezhető.

A kriminalisztikai vagy törvényszéki esetekben a helyzet lényegesen bonyolultabb és bizonytalanabb. A beszélő személy ismeretlen, következésképpen nincsen „tárolt” beszédminta. Jó esetben számíthatunk, ha van gyanúsított vagy gyanúsítottak, ez kiindulást jelenthet a személyazonosításhoz. A feltételezett eredeti beszélőnek azonban ekkor nem célja, hogy természetesen, tisztán, megfelelő hangerővel beszéljen; az akusztikai-fonetikai összevetés tehát nehezedik. Mintegy 15%-ra tehető ezekben az esetekben, hogy a beszélő akaratlagosan megváltoztatja a beszédét (Künzel 1995). A leggyakoribb ilyen torzítások a sutogás, a megemelt hangfekvés és a zárt szájjal képzett beszéd. A rögzített beszéd rendszerint zajos, egy szűk frekvenciatartományban jelentkezik, a hasznos paraméterek tehát erősen csökkentett számban

vannak jelen (nemritkán csak 15-30 másodpercnyi anyag áll az elemző rendelkezésére).

A leglényegesebb különbség a kétféle beszélőazonosítás között a lehetséges beszélők számának különbsége. Az egyik esetben tulajdonképpen a beszélő személyének igazolása történik meg; a kriminalisztikai esetekben pedig valóságos azonosítás a cél. A beszélő azonosításához rendszerint háromféle megközelítésmódot használnak:

- (i) hallás alapú elemzések (általában képzett szakemberek, elsősorban fonetikusok részvételével),
- (ii) akusztikai-fonetikai analízis széles sávú spektrogramok alapján,
- (iii) félautomatikus, speciálisan fejlesztett számítógépes elemző rendszerek alkalmazása.

A hallás alapján történő azonosítás tulajdonképpen percepciók teszt sorozat, amikor a hallgató a rögzített beszédet igyekszik a feltételezett személlyel azonosítani (a hallgató emlékezetében tárolt minta alapján). A beszélő személyt nem ismerő lehallgatók a feltételezett egyezéseket próbálják meghatározni a rövid idejű memóriában tárolt beszédminták összevetésével. Mindkét esetben előfordulhat olyan feladat is, amikor – kizárásos alapon – azt kell megmondani, hogy melyik az a beszélő, aki biztosan nem azonosítható az eredetivel. A szakemberek olyan kérdésekre is tudnak valószínű választ adni, mint a nyelvjárás lehetősége, beszédhiba, a szociális háttér, iskolázottság, becsléssel az életkor, a beszédbeli jártasság. A fonetikus és nem fonetikus hallgatók beszélőazonosítási eredménye között nagy különbség is lehet. Köster (1987) azt találta kísérletében, hogy míg a fonetikusok 100%-ot értek el, addig a nem-fonetikusok csak 89%, 94%-ot. Kollektív beszédének 30 másodperces részletei elegendőek voltak ahhoz, hogy tökéletesen azonosítható legyen a beszéd (Ladefoged 1978).

### **A beszélő felismerésének humán képessége**

Az anyanyelv-elsajátítás folyamán kialakulnak azok a neurális spektrogramok az agyban, amelyek lehetővé teszik, hogy a gyermek a beszélő artikulációs sajátosságaitól függetlenül képes legyen a beszédhangokat azonosítani, a szavakat felismerni. Nem tudjuk még pontosan, hogy vajon ezek a neurális spektrogramok – mint ahogy



megnevezésük felveti – valóban hasonlatosak a beszédről készült akusztikai regisztrátumokkal, a spektrogramokkal. Annál is inkább, mivel a spektrogramok mindig egyediek, a neurális spektrogramok pedig szükségszerűen valamiféle általánosított képek kell, hogy legyenek. Feltételezhetően a hangsor(ok)ra szignifikánsan jellemző invariáns jegyeket tartalmaznak, amelyek egyúttal információval szolgálnak a beszélő személyére vonatkozóan is. Minél hosszabb az ugyanazon beszélőtől származó szöveg, a hallgató annál biztosabban képes a beszélőt felismerni. Ennek alapján az is feltételezhető, hogy a beszéd hallgatásakor aktiválódó neurális spektrogram-sorozatban valamiképpen hangsúlyozottabbá válnak a beszélőt azonosító paraméterek. Ezek a feltételezések vezettek a matematikai megoldások kereséséhez, amelyek azonban nem hozták meg a várt eredményt. Pontosan az a valami hiányzott, ami a legteljesebben emberi; az emberi agynak az a képessége, amelyet emlékezésnek nevezünk.

Az emlékezést úgy határozhatjuk meg, hogy olyan folyamat, amelynek során régebben észlelt tárgyak, jelenségek és események kép/pékeit és ezek összefüggéseit idézzük fel anélkül, hogy az azokat létrehozó ingerek vagy ingeregységek éppen hatnának ránk. Az emlékezés az objektív valóságnak a tudatban történő visszatükröződése. Az emlékképek a múltbeli észlelések, élmények reprodukciói. A beszélő személy felismerésére vonatkoztatva két dolog alapvetően fontos: szükséges a megfelelő inger, valamint a felidézés képessége. Az észleletek, feldolgozott ingerek megjegyzéséhez az szükséges, hogy létrejöjjön az emlékenyom, amely az ismétlések során bevészódik. Minél gyakoribb az ismétlődés, annál nagyobb mértékű a bevészódás. Ha ritkán hallunk valakit beszélni, lassabban, nehezebben azonosítjuk a beszédet a beszélővel. Minél gyakoribb a beszéd akusztikai élménye, annál gyorsabb és biztosabb lesz a beszélő személy felismerése.

Az emlékenyomok felidézése többféleképpen történhet, általában valamiféle asszociáció révén. A felidézés alapja az a kapcsolat, amely bizonyos fokig már a bevészódéskor jelen van. Az asszociáció az emlékezésben azt jelenti, hogy a kialakult szinoptikus kapcsolatok működése révén az egyik emlékenyom aktiválása egy vagy több hozzá kapcsolódó emlékenyomot is aktivál. A beszédre vonatkozóan általános összefüggések is megfogalmazhatók. Nem véletlen például az alap-

hangmagasság és a testalkat vagy a hangszínezet és az arcforma kapcsolata; ezek tudományos igényű kutatásáról azonban még alig beszélhetünk. A hallgatók asszociációs képessége a beszéd és a beszélő személyének felidézését illetően különböző. Vannak, akiknél gyorsan történik a bevéődés, gyors a megfelelő neurális spektrogram aktiválása és ennek következtében a beszélő felismerése. Másoknál ezek a folyamatok lényegesen lassabban alakulnak ki, illetőleg mennek végbe.

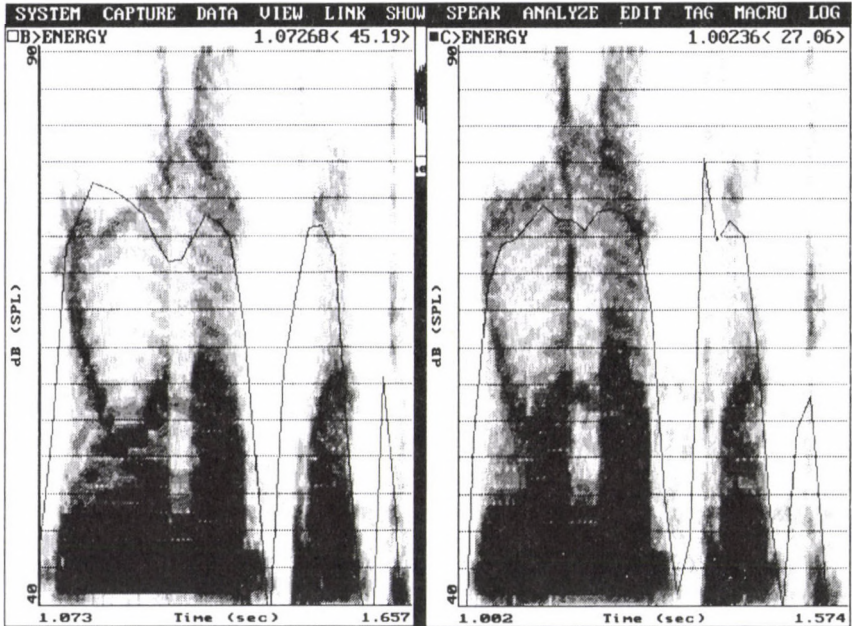
### **A beszélő személy felismerésének tényezői**

A fentiekben az alapvető feltételt – a beszélő személy ismertségének megfelelő szintjét – már tárgyaltuk. A következő, egzaktan nehezen megfogható, ám a pszicholingvisztikában jól ismert tényezőt vesszük számba, az ‘elvárás’ faktorát. Saját elvárásaink hatással vannak a beszélő személy sikeres felismerésére. Ha egy jól ismert személynek telefonálunk, rövid ideig egy hozzá hasonló hangú beszélőt is elfogadunk a kívánt beszélőül az elvárás miatt. Ha várjuk valakinek a hívását, azonnal felismerjük, ha megszólal. Ugyanennek a beszélőnek az azonosítása nehezebb, ha nem feltételeztük, hogy éppen az a személy fog telefonálni (Ladefoged 1978).

A beszélőre jellemző neurális spektrogram nyilvánvalóan tartalmazza mindazokat a nyelvi/beszédbeli tényezőket, amelyek alapján azonosítjuk a személyt. Amennyiben ezt nem kérdőjelezzük meg, akkor valójában mi okozza az egyénre jellemző akusztikai tulajdonságok műszeres kimutatásának nehézségét? Elsősorban az, hogy a beszédinformációt továbbító akusztikus rezgések a hangképző rendszer tehetetlensége következtében kvázistacionárius jellegűek. Ez azt jelenti, hogy a rezgések paraméterei általában korlátozott ideig tekinthetők állandónak. Az előbbiekből következik, hogy a beszéd közben létrehozott hangsorok nem ismételhetők meg még egyszer teljesen azonosan. Az 1. ábrán a „*Jó napot*” hangsor spektrogramja és hangsoron belüli intenzitásviszonyai láthatók ugyanazon személy ejtésében 1 nap eltéréssel. A lehető legjobb, torzításmentes megjelenítés érdekében a hangsort 50000 minta/s-os mintavételezési sebességgel digitalizáltuk és Hamming ablakfüggvényű 71 Hz-es szűrővel analizáltuk. Az ábra bal és jobb oldalának vizuális összehasonlítása alapján is megállapítható, hogy az időben később készült hangfelvételtől regisztrátumán



(jobb oldali) a formánsok és az intenzitás értékei lényeges eltérést mutatnak a korábbi felvételtől készült regisztrátumhoz (bal oldali) képest.



1. ábra

A „Jó napot” hangsor spektrogramja 0-3 kHz-es tartományban

A hatvanas, hetvenes évek nem túlzottan széleskörű kutatásai a beszédhangok akusztikai szerkezetében jelölték meg a meghatározó paramétereket. Elsősorban a magánhangzók harmadik formánsát gondolták ebből a szempontból jelentősnek, amelyről azóta egyértelműen bebizonyosodott, hogy nem is igazán jellemző és messze nem ele-

gendő az egyén azonosításához. Minden egyes akusztikai paramétert igyekeztek megvizsgálni azzal a nem titkolt céllal, hogy a mindent kizárót vagy legalábbis a legjellemzőbbet megtalálják. Ha azonban csak egy formánszt nézünk is (jelen esetben a harmadikat), akkor is három, numerikusan kifejezhető adattal állunk szemben: a formáns frekvenciaértékével, sávszélességével és intenzitásával. Figyelembe véve azt az egyáltalán nem elhanyagolható tényt, hogy e három összetevő állandó változása a beszéd velejárója, akkor nehéz elméletileg is feltételezni azt a számértéket, amely az egyénre jellemző lehet. Ha pedig nem tudunk meghatározni egy vagy néhány konkrét frekvenciaértéket (maximum  $\pm 30$  Hz eltéréssel), akkor a személyazonosítás számértékek alapján nem valószínűsíthető. Egyelőre még nem vettük figyelembe azt, hogy a formánsok értéke függ a hang hangkörnyezetétől is.

A beszélő felismeréséhez – spektrális elemzéssel – a megoldás a teljes frekvenciatartományban megjelenő beszédhang azonosítása lenne (80 Hz-től 16 000 Hz-ig). Tekintetbe kell venni azonban azt is, hogy a beszédhangok akusztikai elemzésével a beszédhangok határértékeit határozták meg. A beszédészlelési kísérletek eredményei pontosították (minden vizsgált nyelvben) azt, hogy a hallgató személy a kérdéses beszédhangot milyen frekvenciatartományban azonosítja biztosan. Nem egyetlen érték felel meg tehát egy hang valamely formánsának a percepcióban, hanem egy határérték, amely 100-200 Hz-es sávot is felölelhet. A formánsstruktúrában az F1 értéke általában szűkebb, mint az F2-é, míg az F3-é nagyobb az F2-énél. A formánseltolódás értéke a beszélő nemétől is függ. Női ejtésben az F2, F3 jobban változhat, mint férfi ejtésben. Ezek a formánsérték-változások ugyanazon beszélő ejtésében kevésbé térnek el egymástól, viszont azonos hangsorok különböző időpontban történő ejtésekor már 30-100 Hz-es eltérések is fennállhatnak. Létezik olyan kutatási eredmény is (Hollien 1977), amelyik nemcsak az F3 jelentőségét kérdőjelezi meg, hanem az egyéni hangszín akusztikai megfelelőjét a telefonsávon kívülre eső összetevőkben feltételezi.

A hetvenes évek végének kutatási eredményei szerint az alaphangmagasság majdnem elegendő kulcs az egyén hangjainak felismerésére (innentől már csak egy lépés magának, a személynek az azonosítására). Úgy tűnik azonban, hogy a pozitív eredménnyel zárult megkülön-



böztetési kísérletek háttérében inkább a hallgatók jól működő rövid idejű memóriája állt, semmint az F0 mint egyértelmű felismerési tényező (Doehring–Ross 1971). Azt gondolták, hogy a vokális traktus fontosabb a beszélő azonosításában, mint a larynx-forrás (vö. Hecker 1971). Ezek a laryngográfiás kísérletek is sikerrel zárultak; ismert személyek közül egy mondat alapján azonosították a kérdéses személyt. Valamennyi beszélő felismerése csak az F0 alapján azonban csak 60–70%-os eredményt hozott.

Az akusztikai elemzések döntően a spektrográfián alapszanak; a következő paramétereket vizsgálják (különböző nyelvekben): formáns-sávszélesség, központi formánsfrekvenciák, maximumpontok, a rés- és zárhangok zörejfrequenciái, átmenetek és még valami, amit úgy neveznek, hogy 'sajátos spektrográfiás alakzat', de közelebbről nem meghatározható paraméter (Künzel 1995). Tekintetbe veendők még a beszédtempó, illetőleg az artikulációs sebesség, a hezitációs jelenségek és a dallammenet. A kutatók azonban egyetértenek abban, hogy a spektrogramok elemzése nem nyújt egyértelmű kulcsot a beszélő személy felismeréséhez. Az alapvető kiindulás mégis a beszéd akusztikuma. A Los Angelesben kifejlesztett beszélőazonosító rendszer (Nakasone–Melvin 1988) például 14 paramétert használ (az időtől a spektrumig). Ezzel a rendszerrel állítólag 98%-os pontosságot lehet elérni (a kísérletek 50 férfi beszélőtől származó beszédmintát tartalmazó adatbázison folytak).

A Hollien és munkatársai által kifejlesztett fonetikai alapú rendszer (SAUSI) olyan paramétereket használ az azonosításhoz, mint az F0, a csendes szünetek száma és hossza, a beszédtempó vagy a magánhangzók időtartama (Hollien 1990).

Több kísérletsorozatban vizsgálták a jelentés szerepét a beszélő felismerésében. Nem a nyelvi, stilisztikai sajátosságok tekintetében, a kérdés csupán az volt, hogy a szöveg érthetősége összefügg-e a beszélő személyének felismerésével. Az eredmények azt mutatták, hogy nem, az a beszélő azonosítása független a beszéd szemantikai sajátosságaitól (Janota 1967; Lariviere 1972; Schlichting–Sullivan 1998).

A leírtakból látható, hogy meglehetősen eltérőek a vélemények abban a tekintetben, hogy melyik a beszédnek az az összetevője, amelyik egyértelmű azonosítást tesz lehetővé. Az alaphangmagasság értéke, a

formánsfrekvenciák, a beszédhang mikrintonációs szerkezete, a beszédhangok egymáshoz viszonyított intenzitása, a beszéd időszerkezete mind-mind olyan paraméterek, amelyeket újra és újra meg kell vizsgálni az egyéni hangszínezet szempontjából. Azt vagy azokat a paramétereket kell megtalálnunk, amelyek mind a szegmentális, mind a szupraszegmentális szerkezetet tekintve, a legkisebb értékkel változnak, azaz közel állandó jelleggel reprezentálják a beszélő személy beszédét.

Láttuk, hogy nagy jelentosege van a beszédminták minőségének. A fizikai értelemben jó minőséggel rögzített minták összehasonlítását a beszéd teljes spektrumában el lehet végezni. Megemlítendő az az eset, amikor az összehasonlítandó hangfelvételek rossz jel/zaj viszonyúak, és a kérdéses felvétel nem egységes telefonhálózaton belül készült. A minőségen kívül fontos szerepe van a minták időtartamának, az egységnyi időtartam alatt elhangzó információnak, valamint a szöveg spontaneitásának.

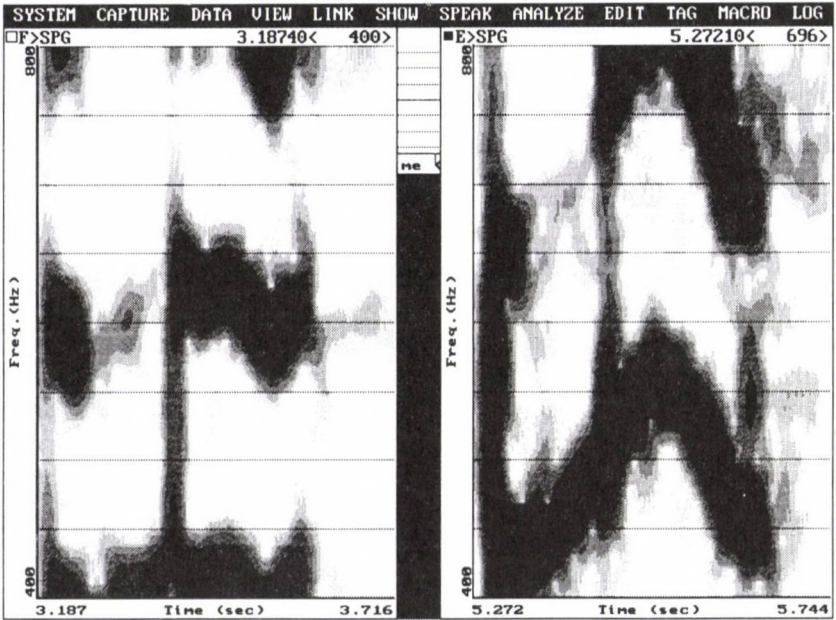
### **Módszertani megközelítések**

Elsődleges célunk az volt, hogy a már korábban említett, ún. „sajátos spektrográfias alakzat” további tanulmányozásával, illetve kísérleti igazolásával olyan vizsgálati módszert dolgozzunk ki, amelynek segítségével 1. a korlátozott (rövid) időtartamú, rossz jel/zaj viszonyal rendelkező beszéd vizuálisan és numerikusan is összehasonlító és 2. a kapott eredmények alapján az azonosítás nagy biztonsággal elvégezhető.

Az MTA Nyelvtudományi Intézetének Fonetikai Laboratóriumában olyan eljárást dolgoztunk ki, amely lehetővé teszi rossz technikai körülmények között rögzített hangfelvételeknél a személyazonosítást (azonos hangsorok esetében). A jelen tanulmányban ismertetendő elemző eljárásunk lényege az, hogy a hang spektrumából csak a felharmonikusok vizsgálatával foglalkozunk, függetlenül attól, hogy az adott tartományban van-e formáns vagy nincs. Mivel a felharmonikusok alacsonyabb frekvenciatartományban intenzívebben vannak jelen, ezért alkalmasabbak műszeres elemzésre. A 400 Hz és 800 Hz közötti tartományban lévő felhangok dinamikus változását és elhelyezkedésének numerikus értékét vesszük figyelembe. Megfelelően választott



hangszelet (frame), illetőleg sáv szélesség esetében a szegmentált részből készített hangszinkép alapján az összehasonlítás elvégezhető. A „*hogy ma este*” hangsor „*maes*” részletéről készült regisztrátumok illusztrálják a fentieket a 2. és 3. ábrán.

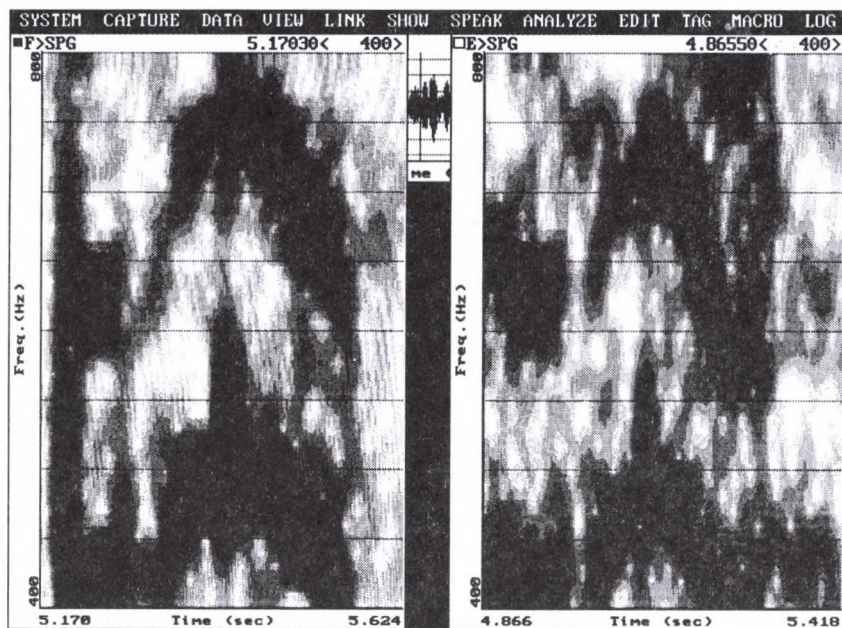


2. ábra

A *hogy maes* hangsor ejtése RM jelű (bal oldal) és SZM jelű személy (jobb oldal) esetében

A példában a személyazonosítás telefonvonalról, rossz minőségben rögzített hangfelvételtől szegmentált közlésrészlet alapján történt. A kiválasztott rész időtartama 560 ms. A lehetséges beszélő személyektől az előre meghatározott szöveget stúdióminőségű felvételen rögzítettük, és a négy hangsorral készített spektrumot vizuálisan

elemeztük. Az RM jelű spektrogramról elmondhatjuk, hogy a felhangstruktúra dinamikus változása semmiképpen nem azonos az etalonfelvétellel. Az SZM jelű személy hangjáról készült spektrogram már mutat némi hasonlóságot, de numerikus értékei nem azonosak. A PH jelű személy azonban minden tekintetben azonosnak mondható a telefonról rögzített személy hangjával. (Megjegyezzük, hogy a vizsgálatban valójában 9 személy vett részt, de a többi részvevő hangszíneke lényeges eltérést mutatott az eredetihez képest, ezért ezekre itt nem térünk ki).



3. ábra

A *hogy maes* hangsor ejtése PH jelű személytől (bal oldal) és a telefonról rögzített hangfelvétel

E módszer hátránya, hogy csak azonos hangsorok esetében alkalmazható. A szemléletesség kedvéért közöljük a hangsorok teljes spektrumát 5 kHz terjedelemig (4a, 4b. ábra). Az ábrákból jól látható, hogy az azonosítás ilyen megjelenítési formában nem végezhető el.

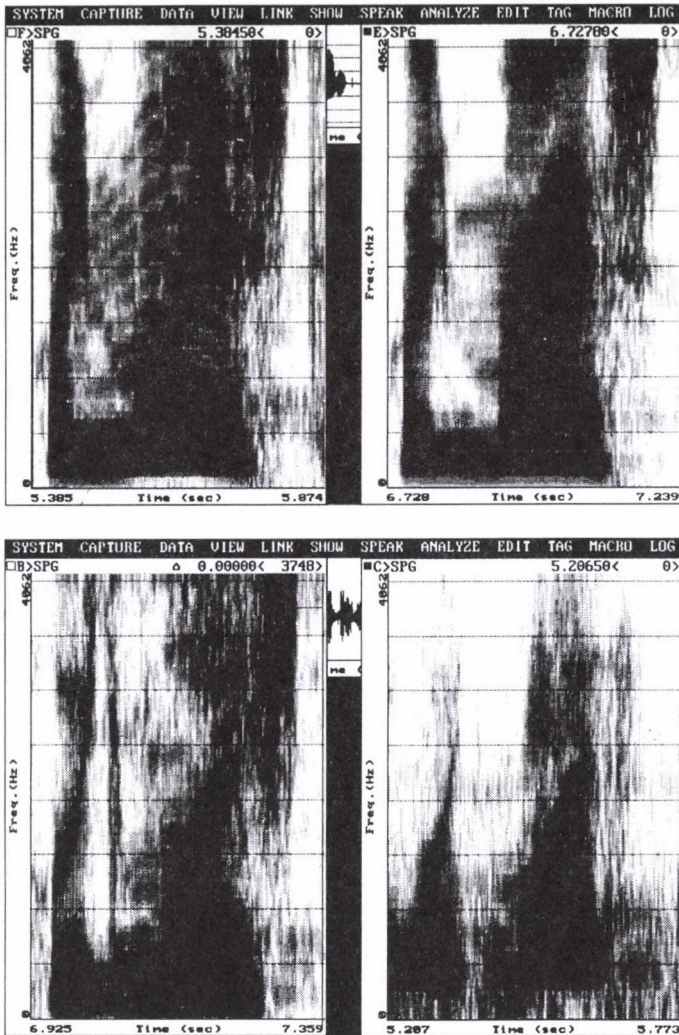
A felhangok változásának elemzésén alapuló eljárás magában hordozza a hibalehetőséget is. Abban az esetben, ha a rögzített szöveg – rövidsége miatt – nem tartalmaz hasonló beszédhangokat, hangkapcsolatokat, ez az azonosítást erősen megnehezíti. Ezért következő lépésként olyan módszerrel kísérleteztünk, ahol nem a beszélő személy által ejtett hangsorokat tekintjük kiindulási alapnak, hanem annak kisebb egységét, magát a hangot. A hosszan ejtett magánhangzó, illetve a zöngés mássalhangzó ugyanis bizonyos fokig magában foglalja a személy egyéni hangszínezetét is.

Feltételezésünk szerint a felhangoknak is hordozniuk kell a beszélő személyre jellemző egyéni sajátosságokat. A formánsstruktúrát leválasztva tehát, olyan felharmonikusokat keresünk, amelyek a legkevésbé esnek egybe (azaz megfelelő távolságban vannak) a formáns helyével. A formáns helytől távol levő felhang ugyanis a kívánt felbontással elemezve, magában hordozza a hangszalagrezgés egy teljes periódusában bekövetkezett változást. Ez az eltérés pedig spektrálisan megjeleníthető. Az így megjelenített, megfelelő számú spektrum összehasonlításával, kialakítható egy olyan analízáló stratégia, amely rövid idő alatt nyújt értékelhető adatot, és jól reprezentálja a személy hangjának bizonyos sajátosságait.

Mindezek igazolására kísérletet végeztünk, ahol a beszédminta formáns- és felhangstruktúráját vizsgáltuk. A kiválasztott „tisztá” felhangot a jó melléknyaláb-csillapítás érdekében Hamming ablakfüggvényű keskeny sávú digitális szűréssel megtisztítva (a többi felhangtól), vetjük alá az összehasonlító spektrális vizsgálatnak. A cél annak megítélése, hogy az így kiválasztott felhang spektruma milyen képet ad a hang teljes időtartamában, illetve, hogy elegendő-e 8-10 periódus a vizsgálat céljára.

A kísérlet első fázisában 5 beszélővel végeztük el a vizsgálatokat úgy, hogy a kísérleti személyektől rendelkezünk 23 évvel ezelőtt készített hangfelvételekkel.





4b. ábra

Az SZM és az RM jelű személyek ejtéséről készült teljes spektrum (fent); a PH jelű személy hangjáról (lent, bal oldal) és a telefonról készített hangfelvétel teljes spektruma (lent, jobb oldal)



Az öt személy hanganyagát CSL 4300B típusú digitális jelfeldolgozóval vizsgáltuk. A vizsgálat menete a következő volt:

1. A kiválasztott szöveget digitálisan rögzítettük. A rögzítés mintavételezési sebessége 50000 minta/s.

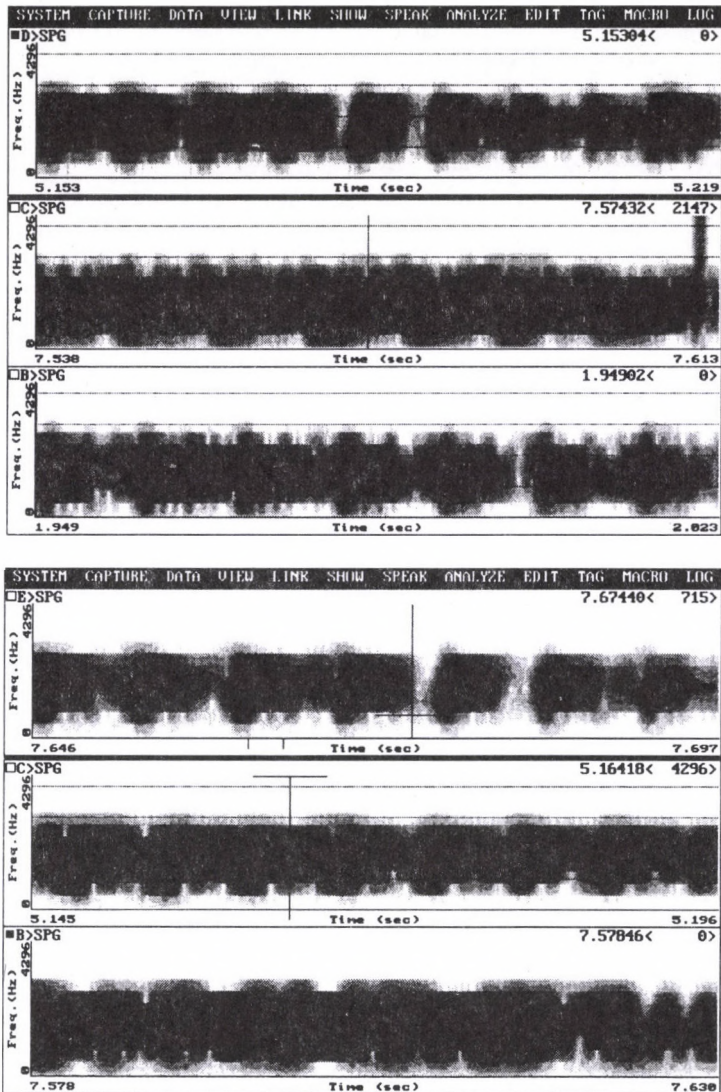
2. A bevételezett hanganyagból a kiválasztott hangot (ebben az esetben ez az [l] mássalhangzó volt) szegmentáltuk.

3. A megfelelő hosszúságú mintát (kb. 70 ms) keskeny sávú Hamming ablakfüggvénnyel 300 Hz sávszélességben szűrtük, majd az intenzitását többszörösen megnöveltük.

4. A keletkezett mintáról keskeny sávú spektrogramot készítettünk. A spektrogramok összehasonlítása a következő eredményhez vezetett. A különböző személyektől származó zöngés mássalhangzók periódusonként más-más elhelyezkedésű intenzitásmaximumot mutatnak. A 6. ábrán látható, hogy a felső részben lévő spektrogramon a periódusok gömbszerű alakot vesznek fel, lefelé mutató nyúlvánnyal, a középső részben elnyújtott formát láthatunk, felfelé mutató nyúlványokkal. Az alsó részben látható forma hasonlít ugyan kissé a felsőhöz, de a góccok maximumpontjai inkább felfelé mutatnak. Az ábra 4. sorában az első sorban lévő személy hangja ismétlődik meg, de a két ejtés között 1 hét különbség van. Az 5. sorban, a kissé ellaposodó periódustól eltekintve, nagyjából azonos jelek láthatók. A 6. sor viszont periódusonként eltérő képet mutat lefelé mutató nyúlványokkal.

### **Következtetések**

A beszélő személy azonosítása a beszéde alapján már bizonyos múltra tekinthet vissza a magyar szakirodalomban (Gordos–Takács 1983; Gósy 1996; Nikléczy 1996;), de rendszeres akusztikai-fonetikai és percepció vizsgálat a két éve indult meg. A munkálatok részlegesen ugyan támaszkodhatnak a nemzetközi szakirodalomban leírt eredményekre, azonban a nyelvspecifikusság ténye mindig új feladat elé állítja a kutatót. Az alábbiakban összegezzük azokat a megállapításokat, amelyek részben elméleti meggondolás, részben gyakorlati tapasztalat, illetőleg saját kísérleti eredményeink alapján már egyértelműen megfogalmazhatók. Ezek a megfogalmazható kijelentések nem egyszer sok-sok órá elemző munkán, számtalan adat sokféle feldolgozásán alapulnak (magukban foglalva a kutatás zsákutcait is).



6b. ábra

Az [l] hang átlagosan nyolc periódusáról készült spektrografikus kép  
különböző személyek ejtésében

1. A beszéd akusztikuma jellemző a beszélőre, oly mértékben, hogy az akusztikai-fonetikai paraméterek alapján a beszélő azonosíthatóvá válik.
2. Az elméleti megállapítást a humán beszélőfelismerő képességünk is alátámasztja.
3. A beszéd alapján történő közel-objektív személyazonosítás számátalan tényező függvénye. Ezek részben külső faktorok (pl. a beszédrengztési körülmények) és belsőnek tekinthetők (pl. a beszélő kooperációs készsége).
4. Többféle eljárás is célravezető lehet; az adott módszert mindig a beszélőfelismerés célja, a beszédminta és egyéb körülmények határozzák meg.
5. Pillanatnyilag nincs tudományos válasz arra vonatkozóan, hogy hány vagy mely paraméterek azok, amelyek az egyén felismerését kétséget kizáróan biztosítják. Nem zárható ki az, hogy az agyban tárolt neurális spektrogram aktiválása egészen különböző azoktól az akusztikai eljárásoktól, amelyek révén a beszéd egyéni jegyeit igyekszünk meghatározni.
6. Különböző aspektusú akusztikai-fonetikai és percepciós kísérletek (és rengeteg adatfeldolgozás, -tárolás és összegzés) szükségesek ahhoz, hogy a probléma megoldásához közelebb jussunk.

A továbbiakban a leírt eredmények alapján tervezzük a vizsgálatainkat és kísérleteinket.

### Irodalom

Doddington, G.R.—Helms, R.E.—Hydrick, B.M.: Speaker verification III. Texas Instruments Inc. Report for RDAC, Rome, New York 1976.

Doehring, D.G.—Ross, R.W.: Voice recognition by matching to sample. J. of Psycholinguistic Res. 1. 1972, 233-142.

Gordos Géza—Takács György: Digitális beszédfeldolgozás. Műszaki Könyvkiadó. Budapest 1983.

Gósy Mária: A beszéd akusztikai szerkezetének állandóságáról. In: Nyelv, nyelvész, társadalom. Emlékkönyv Szépe György 65. Születésnapjára barátaitól, kollégáitól, tanítványaitól. II. Szerk.: Terts István. Keraban Könyvkiadó. JPTE. Pécs 1996, 66-75.

Hecker, M.: Speaker recognition: an interpretative survey of the literature. A.S.H.A. Monogr. 16. Washington, D.C. 1971.



Hollien, H.: Speaker identification by long-term spectra under normal and distorted speech conditions. *JASA* 62. 1977, 975-980.

Hollien, H.: *The Acoustics of Crime*. Plenum Press. New York, London 1990.

Janota, P.: Personal characteristics of speech. *Trans. Of the Czechoslovak Academy of Sciences – Social Sciences Series* 77/1. 1967.

Künzel, H.J.: Field procedures in forensic speaker recognition. In: Windsor Lewis, J.: *Studies in General and English Phonetics. Essays in Honour of Professor J. D. O'Connor*. Routledge. London 1995, 68-85.

Ladefoged, P.: Expectation affects identification by listening. *Language and Speech* 21/4. 1978, 373-375.

Lariviere, C.: Acoustic and perceptual correlates to aural speaker identification. In: Rigault, A. (ed.): *Proc. 7<sup>th</sup> ICPHS*. The Hague 1972, 558-564.

Nakasone, H.–Melvin, C.: Computer assisted voice identification system. *Proceedings IEEE-ASSP*. 1988, 587-590.

Nikléczy Péter: Beszélő személy azonosítása szűk frekvenciás szavak alapján. In: *Beszéd kutatás '96*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1996, 20-31.

Schlichting, F.–Sullivan, K.P.H.: Can voice imitation be detected in voice line-ups in a language unknown by the listeners? *Phonum* 6. 1998, 105-118.

Stevens, K.N.: *Acoustic Phonetics*. MIT Press. Cambridge, Mass. 1998.

A kutatás a T 025965 sz. OTKA-munkálat keretében folyt

# EGYÉNI KÜLÖNBSEGEK AZ ARTIKULÁCIÓS TEMPÓ PERCEPCIÓJÁBAN

Gocsál Ákos

Janus Pannonius Tudományegyetem, Tanárképző Intézet

## Bevezetés

A hangjelenségek egyik legfontosabb tulajdonsága az időbeliség. A beszéd esetén ez nem pusztán azt jelenti, hogy egy adott hang meghatározott ideig szól, hanem arra is utal, hogy a kibocsátott hang frekvencia-összetétele és intenzitása időben változik; így jönnek létre a különböző beszédhangok. Az utóbbi évtizedekben számos hazai és külföldi kutatás foglalkozott a beszéd időszerkezetével. Az objektív mérések mellett azonban számos igen érdekes kérdést vet fel a beszéd időviszonyainak észlelése. A jelen dolgozatban egy beszédtempó-percepciói kísérlet eredményeit ismertetjük néhány lehetséges magyarázattal.

## Az artikulációs tempó változásainak fonetikai jelentőségéről

Az artikulációs tempó változtatása számos következményt von maga után. Korábbi kutatások arra engednek következtetni, hogy a gyors beszédnek igen sajátos fonetikája és fonológiája van. Siptár (1991) például azt mutatta ki, hogy a felgyorsult beszéd esetén jellegzetes, a normál tempójú beszédétől eltérő fonológiai szabályok (hangok kiesése, asszimilációja, rövidülése) érvényesülnek.

Az artikulációs tempó változása jelentést is hordozhat magában. Beszédünk lassításával és gyorsításával – tudattalanul vagy szándékosan – érzelmeket is ki tudunk fejezni. Scherer (vö. Scherer 1995) az ide vonatkozó szakirodalom összegzése kapcsán megállapította, hogy például öröm és harag kifejezésekor gyakori a gyorsabb beszéd, szomorúság esetén viszont általában lassabban beszélünk. De a gyorsítást és a lassítást egyéb fonetikai eszközökkel, például megfelelő szünettartással kombinálva, hatásosan tudunk egyes szövegrészeket kiemelni, nyomatékosítani is.

A szervezet öregedése, fáradása is megjelenik a beszédtempóban: Balázs (1993) két ismert színész, Dajka Margit és Páger Antal fiatalkori és időskori hangmintáit összehasonlítva azt találta, hogy egy perc alatt Dajka 20, Páger pedig 40 szóval mondott többet a fiatalkori felvételeken. (Az azonban, hogy a beszédtempó értéke milyen mértékben játszik szerepet a beszélő személy életkorának megbecslésében – kizárólag a hangja alapján –, további vizsgálatok tárgyát képezi.)

A beszédtempó-percepciók kísérletek közül Gósy (1991) néhány eredményéről kell említést tennünk. Egyrészt: minimális, de jól meghatározható különbség mutatkozik a különböző tempóval beszélő egyének tempóérzékelésében. Másrészt: a szélsőségesen „gyors”, illetve „lassú” beszédű egyének tempóérzékelése sok szempontból hasonló, szemben a „normál” tempóval beszélőkével. A beszédértést tekintve is kimutathatók különbségek, a gyorsabban beszélők beszédértése szignifikánsan jobb volt, mint a lassúaké, de ők több hibát is vétettek a beszédértés során.

A fentiek alapján arra következtethetünk, hogy a beszédtempó számos egyszerre jelentkező, de egymástól nagyban eltérő tényező eredőjeként jelentkezik (ilyen tényezők az életkor, személyiség, érzelmi állapot, az elmondott szöveghez való viszonyulás, beszédtechnikai jártasság, a hallgató személye stb.). Feltehető tehát, hogy a beszédtempók igen nagy változatosságot fognak mutatni, még akkor is, ha a beszédprodukció az egyes személyek esetén ugyanazon körülmények között és ugyanazon kontextusban történik. De vajon jelentkezik-e különbség a beszédtempó észlelésében a különböző személyek esetén? Erre a kérdésre keressük a választ a jelen kísérlettel.

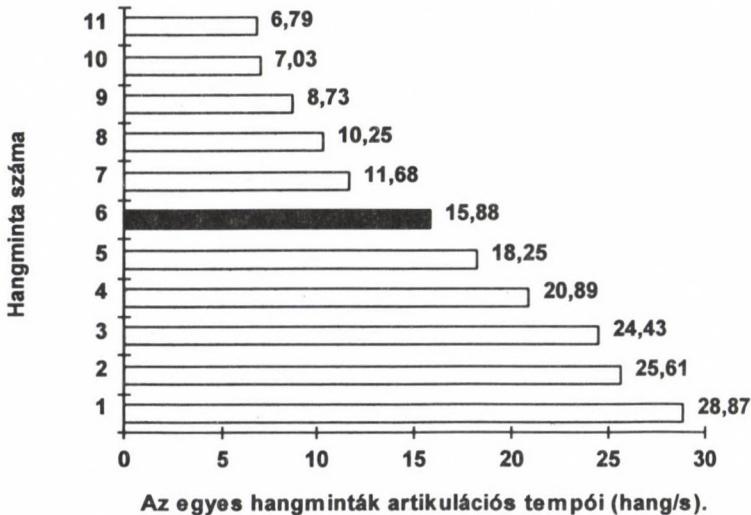
### **Hipotézisek**

Korábbi kísérletünk során (Gocsál 1998) utaltunk arra, hogy ugyanazon fonetikai jelenség – az adott vizsgálatban a beszélő személy életkorának megbecslése hangja alapján – percepciója szempontjából a hallgatók legalább négy fő csoportra oszthatók. Vannak, akik jó „megítéléssel” rendelkeznek, így általában a valós életkorhoz közeli becsléseket adnak. Mások alul-, megint mások pedig inkább felülbecslik a beszélő személy korát. A negyedik csoportba pedig a rendszerte-



len becsléseket adók tartoznak, náluk a megadott életkori értékek gyakorlatilag véletlenszerű eloszlást mutattak.

Hipotetikusan azt állítjuk, hogy az itt vázolt percepció viselkedési típusok más fonetikai jelenségek esetén is kimutathatók, természetesen az adott jelenség sajátosságaira vonatkoztatva. A jelen kísérletben az artikulációs tempó észlelésével foglalkozunk, így tehát azt várjuk, hogy a) lesznek hallgatók, akik a normál artikulációs tempóval lejátszott szöveget gyorsnak fogják ítélni, azaz az ő számukra a normál tempó az irodalmi értéknél lassabb, b) ennek fordítottjára is számíthatunk, tehát olyan hallgatókra, akik a normál beszédtempót lassúnak ítélik, továbbá c) „jó megítélésű” hallgatókat, és d) kérdéses eseteket is várunk, amelyeknél valamilyen szokatlan jelenség mutatkozik.



1. ábra

A percepció kísérletben felhasznált hangminták időviszonyai  
(az eredeti hangmintát sötétebb színnel jelöltük)

### Anyag és módszer

A percepció kísérlet eszközéül egy 2,33 s hosszúságú mondat szolgált („*Kérek szépen egy pohár vizet, mert szomjas vagyok.*”) egy középkorú férfi beszélő ejtésében. A mondatot az MTA Fonetikai Laboratóriumában, Varispeech készülékkel 5 fokozatban felgyorsítottuk, majd 5 fokozatban lelassítottuk. A gyorsítás és a lassítás során a beszédminták frekvenciaviszonyai nem változtak (nem lett mélyebb vagy magasabb a beszélő hangja). Az 1. ábrán megfigyelhető az egyes beszédminták artikulációs tempó szerinti elrendezése, értelemszerűen a kisebb értékek lassítást, a nagyobbak gyorsítást jelentenek. Az eredeti hangminta a 6. számú. Az összesen 11 hangmintát véletlenszerű sorrendben kazettán rögzítettük, majd nyugodt körülmények között lejátszottuk 37 egyetemi hallgatónak. A hallgatók kérdőíveket töltöttek ki, amelyeken az egyes beszédminták tempójára vonatkozóan az alábbi megállapítások valamelyikét jelölhették meg, ötfokú skálán:

A hallott beszédminta...

1. túl lassú ahhoz, hogy hétköznapi körülmények között előforduljon;
2. lassú, de esetenként előfordulhat;
3. hétköznapi beszédhelyzetekben előfordul;
4. gyors, de esetenként előfordulhat;
5. túl gyors ahhoz, hogy hétköznapi beszédhelyzetekben előforduljon.

A hallgatókat megkértük arra is, hogy jelezzék, megítélésük szerint saját beszédtempójuk milyen. Sajnos, erre a kérdésre csak 14-en adtak választ. Az azonban így is elgondolkodtató, hogy míg közülük egyvalaki ítélte saját beszédét lassúnak, ketten normálisnak, a többiek vélekedései a gyors tempó legkülönbözőbb fokozatait mutatták, a „néha kicsit gyors”-tól a „gyakran rám szólnak, hogy hadarok”-ig.

### A kísérlet eredményei

A mérési eredmények jól alátámasztották hipotézisünket, de bizonyos szempontból ki is egészítették azt. Terjedelmi okok miatt az egyes esetekre vonatkozóan csupán néhány tipikus példát ismertetünk.

1. táblázat: A lassabb tempókat kedvelők

beszélő sorszáma	lassítás					normál tempó					gyorsítás	
26	1	2	2	2	2	3	5	5	5	5	5	5
22	2	1	2	2	2	3	4	4	5	5	5	5
33	1	2	3	2	3	4	4	4	5	5	5	5
21	1	1	1	2	2	3	4	5	5	5	5	5

2. táblázat: A gyorsabb tempókat kedvelők

beszélő sorszáma	lassítás					normál tempó					gyorsítás	
12	1	1	1	1	1	3	3	4	4	5	5	5
15	1	1	1	2	2	3	4	4	4	5	5	5
4	1	1	1	1	2	3	4	4	5	5	5	5
32	1	1	2	2	2	3	3	3	4	4	4	4

3. táblázat: Az átlagérték körüli tempót kedvelők

beszélő sorszáma	lassítás					normál tempó					gyorsítás	
36	1	1	2	2	2	3	4	4	4	5	5	5
28	1	1	1	2	2	3	4	4	5	5	5	5

4. táblázat: Kérdéses esetek

beszélő sorszáma	lassítás					normál tempó					gyorsítás	
8	1	1	1	1	1	4	5	5	5	5	5	5
13	2	1	2	2	1	2	4	3	4	4	4	4

#### a) A lassabb tempót kedvelők

Az 1. táblázatban szereplő számok az előzőekben ismertetett ötfokú skála szerint adott válaszokat jelentik. A szemléletesség kedvéért a 2-es, 3-as és 4-es válaszokat eltérő háttérszínnel is megjelöltük.

A lassabb tempót kedvelők csoportjában a 26. sz. hallgató tekinthető „tisztá esetnek”. Ő már a minimális mértékű gyorsítást sem fogadja el, ugyanakkor a másodpercenkénti 7,03 hang még elfogadható szá-



mára. Nála „engedékenyebb” a 22-es és a 33-as számú hallgató, ugyanis ők nem zárják ki a kisebb mértékű gyorsítást. A 21-es számú hallgató adatai is hasonló jelleget mutatnak, de az ő esetében az elfogadható tempóértékek intervalluma jóval szűkebb. Összességében a 37 hallgatónak több mint egyharmadánál találtunk ezekhez hasonló eredményt.

#### **b) A gyorsabb tempót kedvelők**

A 2. táblázat adatai alapján megállapítható, hogy a gyorsított tempókat kedvelőknél az előzőhöz hasonló jellegű, percepciós értékeket kapunk. Itt a 12. hallgató tekinthető „tisztá esetnek”. Tipikusak a 14. hallgató válaszai, rajta kívül még öten adtak ugyanilyen adatokat. A 4. hallgató az artikulációs tempó megítélése szempontjából szigorúbb, az általa elfogadhatónak ítélt tempók által meghatározott intervallum igen szűk. A 32. hallgató viszont a lejátszott tizenegy hangmintából csupán a leglassabb kettőt nem tartja elfogadhatónak, a gyorsabb minták iránti preferenciája azonban egyértelmű: még a 20,89 hang/s artikulációs tempót is elképzelhetőnek tartja átlagos hétköznapi beszédhelyzetekben. Az említettekhez hasonló jellegű adatokat szintén a hallgatóknak több mint egyharmada adott.

#### **c) Az átlagérték körüli tempót kedvelők**

A táblázatban szereplő két hallgatón kívül csak néhányan voltak, akiknél a válaszok nem mutattak elcsúszást valamelyik irányba. „Szigorúbb”, illetve „engedékenyebb” hallgatókat azonban itt is találtunk, ahogy az a 3. táblázatból kiderül.

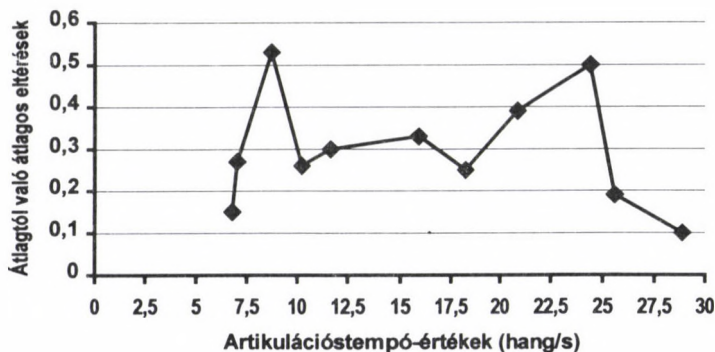
#### **d) Kérdéses esetek**

A 8. és a 13. hallgató annyira sajátos válaszokat adott, hogy érdemes ezeket az előző esetektől külön kezelni (4. táblázat). A 8. hallgató adatai túlzott „szigorra” engednek következtetni: ő teljesen elképzelhetetlennek tartja a középértéktől eltérő tempókat. A 13. hallgató viszont túlzottan „engedékeny”: a kísérletben szereplő leglassabb és leggyorsabb hangmintákat is el tudja fogadni. Ami azonban ennél is meglepőbb, az az, hogy a minimális mértékű lassítást kizárja, a leglassabb tempót azonban elfogadja. Érdekes eredmény még, hogy számára a

normál artikulációs tempót a 20,89 hang/s jelenti – ebből arra következtethetünk, hogy a gyorsabb tempókat kedveli, ám sajátos, hogy a normál tempónak ítélt beszédmintánál lassabban és gyorsabban lejátszott szöveget is egyaránt gyorsnak találta.

### A szórási eredmények

A következőkben azt vizsgáljuk meg, hogy a hallgatóktól kapott adatok átlagtól való átlagos eltérései hogyan alakulnak az egyes hangmintákra vonatkozóan. Az eredményeket a 2. ábra foglalja össze. Minél nagyobb az egyetértés a válaszadók között, annál kisebb a szórás értéke. A grafikon alapján tehát az állapítható meg, hogy a hallgatók igen hasonlóan vélekedtek a két leggyorsabb, illetve leglassabb hangmintáról, azaz ezek az artikulációs tempók véleményük szerint hétköznapi körülmények között nem fordulhatnak elő. Valamivel nagyobb – 0,3 körüli – szórási értékeket kaptunk a kb. 10 és 18 hang/s közötti artikulációs tempóknál. Bizonyos tempóértékek esetén pedig kiemelkedően magas a szórás, ami azt jelenti, hogy ezekben az esetekben nagymértékben megoszlanak a vélemények az adott artikulációs tempó hétköznapi előfordulását illetően.



2. ábra  
Az átlagtól való átlagos eltérések értékei



### **Az eredmények értékelése, lehetséges magyarázatok**

A jelen kísérlet megerősítette, hogy igen nagy különbségek tapasztalhatók ugyanazon beszédminta észlelése során az egyes hallgatóknál. Nem pusztán a gyors/átlagos/lassú artikulációs tempók preferenciáját sikerült kimutatni a különböző egyéneknél, hanem azt is, hogy bizonyos hallgatók a tempóértékek szélesebb, mások szűkebb intervallumát tartják elfogadhatónak. Vajon mi lehet ennek az oka? Egyik lehetséges magyarázatunk az, hogy az egyén hallgatói tapasztalatai mintegy „leülepednek” a tudatban, és így létrejön egy mentális időszerkezet-sztereotípiá. E sztereotípiá határozza meg azt, hogy az egyén mely tempókat miként ítéli meg, méghozzá oly módon, hogy az aktuálisan észlelt és a munkamemóriába került hangmintákat összehasonlíttja ezzel a sztereotíp neurális „időképpel”.

Másik magyarázatunk a beszédfeldolgozási folyamatok sebességével kapcsolatos. Számos olyan pszicholingvisztikai modell ismeretes, amely a beszédészlelés folyamatát értelmezi. A legfontosabb kérdést a szókeresés, és vele összefüggésben a mentális lexikon szerkezete, illetve a szegmentálás problémája veti fel (vö. például Gósy 1998). De bármelyik modellt is alkalmazzuk, figyelembe kell vennünk, hogy a modellek egyes részei közötti információmozgáshoz bizonyos időre van szükség. Másképpen fogalmazva, idő kell a szókereséshez, a jelentés megtalálásához az adott kontextusban, kétértelmű esetben a valószínűbb jelentés mellett való döntéshez, ezzel együtt korábbi ismereteink aktiválásához, asszociációs láncok elindításához és lefuttatásához – hogy csak néhány idevonatkozó mentális műveletet említsünk. Alapfeltevésünk az, hogy ugyanazon beszédesemény észlelése során a beszédértési folyamat egyes műveletei különböző ideig tartanak különböző hallgatók esetén. Léteznek olyan hallgatók, akik ismereteiket, belső sztereotípiáikat gyorsabban tudják aktivizálni, a kívánt szavakhoz is gyorsabban jutnak hozzá. Ebből adódóan, a gyorsabb beszédet is könnyedén fel tudják dolgozni. Amikor lassabb beszédet hallanak, percepció kapacitásuk nincs teljesen kihasználva: ezért inkább a gyorsabb tempókat fogják előnyben részesíteni a lassabbakkal szemben. Hasonló megközelítéssel értelmezhetjük a lassabb tempókat kedvelő egyének percepció viselkedését is. Ők, ha az észlelt szöveg tempója a nekik megfelelőnél gyorsabb, nem tudják teljesen feldolgozni a



hallott beszédet, így inkább a lassabb tempókat fogják előnyben részesíteni. Találunk olyan eseteket is, amikor az objektíven megállapított normál tempó egybeesik az egyén percepció preferenciájával, és a körülötte elfogadhatóként megjelölt tempóértékek szűkebb vagy tágabb intervallumban, szimmetrikusan helyezkednek el. Valószínű, hogy a percepció intervallum tágassága összhangban van az egyén alkalmazkodóképességével.

A szórási grafikon elemzése során az alábbiakra következtethetünk. Láttuk, hogy a kb. 10-18 hang/s tempóértékeket a hallgatók viszonylag kis szórással elfogadták mint hétköznapi helyzetekben (esetleg) elfogadható tempókat. Ezt az intervallumot az adott populáció szigorúan vett tempó-sztereotípiájaként értelmezhetjük. A tempóértékek előfordulását illetően, a legkisebb egyetértés a 0,5 körüli szórásoknál van. Ez azt jelenti, hogy vannak, akik ezeket a tempókat már nem tudják elfogadni, sokak azonban igen. Ennek megfelelően meghatározhatjuk a populáció kevésbé szigorúan vett tempó-sztereotípiáját is, amely a kb. 8,5 - 24,5 hang/s tartományt jelenti. A két intervallum összehasonlításából azt az igen jelentős következtetést vonhatjuk le, hogy bár létezik egy, a hétköznapi beszédhelyzetekre érvényesnek tekintett tempótartományunk, az ettől való eltéréseket különbözőképpen ítéljük meg: **sokkal jobban toleráljuk az ennél gyorsabb beszédet, mint a lassabbat.** Jól megfigyelhető mindez a grafikonon is. A 10-18 hang/s-os tartománytól a lassabb tempók felé távolodva jóval meredekebb a grafikon, mint a gyorsabbak felé. Ez azt jelenti, hogy a maximális szóráshoz tartozó érték, illetve az említett intervallum között a gyorsabb artikulációhoz tartozó szakaszon lényegesen több tempóérték jöhet számításba, mint a lassabbakhoz tartozónál. Ezzel az eredménnyel ismételtelen bizonyíthatónak tűnik, hogy a hallgatók beszédtempója és tempóészlelése szoros összhangban van.

### További kérdések

A kapott eredmények és következtetések számos újabb kérdést is felvetnek. A percepció modellekre visszatérve: bizonyos, hogy bármelyik modellt használjuk is, a modell egyes moduljainak működéséhez időre van szükség. De mennyi időt vesznek igénybe az egyes modulok? Vannak-e a beszédértés folyamatában tipikusan „gyorsan”,

illetve „lassan” végzett mentális műveletek? A működési sebesség szempontjából egész modellekre vonatkozóan is feltehetjük a kérdést: két modell összehasonlítása során például az *A* beszédészlelési modell gyorsabban adja-e ugyanazt az eredményt – azaz a hallott beszéd-nyilatkozás jelentését – mint a *B*?

Gyakorlatiasabb problémákat vet fel az, hogy a tempó változtatása, ahogy a szakirodalmi áttekintés során láttuk, maga is hordozhat jelentést. Kérdés, hogy a beszédtempót különféleképpen észlelő hallgatók a tempó által közvetített, például érzelmi jellegű jelentést hogyan fogják fel, illetve, hogy az életkor becslésben van-e ennek szerepe.

A legérdekesebb kérdést azonban kétségtávol azok az esetek jelentik, amelyek valami miatt nem illeszthetők a lassú/normál/gyors tempót kedvelők rendszerébe. Az említett szokatlanul szigorú, illetve a szokatlanul engedékeny és helyenként meglepően következetlen hallgatók percepcióis viselkedését valószínűleg már extrémnek lehet tekinteni. A szeriális feldolgozást végző modulok részletes vizsgálatával, nagy valószínűséggel megtaláljuk e jelenségek okát. Az itt ismertetett eredmények ugyanis megítélésünk szerint tünetekként értelmezendők, mégpedig bizonyos, ma még sok tekintetben ismeretlen mentális működések megnyilvánulásaiként. A pszicholingvisztikai kísérletekkel pedig – hasonlóan a jelen dolgozatban írottakhoz – éppen ezeket a működéseket szeretnénk feltárni, és még jobban megismerni.

### Irodalom

Balázs Boglárka: Hangképzési zavarok időskorban. In: Beszédkutatás '93. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1993, 156-166.

Gocsál Ákos: Életkorbecslés a beszélő hangja alapján. In: Beszédkutatás '98. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1998, 122-135.

Gósy Mária: A szó felismerése: folyamatok és stratégiák. In: Beszédkutatás '97. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1998, 63-118.

Gósy Mária: The Perception of Tempo. In: Temporal Factors in Speech. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1991, 63-107.

Scherer, P: How Emotion Is Expressed in Speech and Singing. In: Proceedings of the XIIIth International Congress of Phonetic Sciences 2. Ed.: Elenius, K.– Branderud, P. Stockholm 1995, 85-89.

Siptár Péter: Fast Speech Processes in Hungarian. In: Temporal Factors in Speech. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1991, 27-63.

A tanulmány elkészítését a Bács-Kiskun Megyei Fejlesztési Alapítvány és a T 025965 sz. OTKA támogatta.



# SZLÁV ANYANYELVŰEK MAGYAR BESZÉDÉNEK FONETIKAI SAJÁTOSSÁGAI

**Menyhárt Krisztina**  
**MTA Nyelvtudományi Intézete**

## **Bevezetés**

A felnőttkori idegennyelv-tanulás mindig számos nehézséggel jár. Nemcsak a másik nyelv grammatikájának, szintaktikai sajátosságainak és lexikájának megtanulása okoz gondot, hanem az anyanyelvtől gyakran erősen eltérő fonetikai jellegzetességeké is. Utóbbiak hibátlan elsajátítása gyakorlatilag lehetetlen. A jelen dolgozat az akcentusnak nevezett kiejtési nehézségeket igyekszik feltérképezni, mégpedig úgy, hogy objektív mérési adatokkal támasztja alá a szubjektíven (hallás után) tapasztalható eltéréseket.

Az akcentus „azokat a kiejtésbeli sajátosságokat jelöli, amelyek valamely nyelv használatakor annak normatív kiejtéséhez képest eltérés-ként felmerülnek a nyelvközösséghez nem tartozó, de annak nyelvén megnyilatkozó egyén beszédében” (Kassai 1995, 106). Az akcentus elsődleges oka az, hogy az idegen nyelv elsajátításának folyamatában anyanyelvünk artikulációs és percepciós bázisát mozgósítjuk.

Artikulációs bázison a hangképző szervek jellemző mozgásainak, helyzeteinek összességét értjük, amelyekkel a nyelvi rendszer elemeit a beszédben megvalósítjuk (Kassai 1995, 106). Kialakulása alapvetően a percepció révén, auditív úton, utánzással történik. A percepciós bázis az a sajátos hallószervi és idegrendszeri mechanizmus, amelynek révén a levegőrezgések formájában továbbított beszédjeleket felfogjuk, és kibontjuk a bennük foglalt tartalmat (Gósy 1989). A két bázis az anyanyelv-elsajátítás folyamatában épül ki, és ezzel összefüggésben egyre szűkül az ősbázis, és egyre inkább csak az anyanyelvünkre jellemző beszédjelenségeket képezzük. Éppen ez az oka annak, hogy minél közelebb kerülünk a felnőttkorhoz, annál nehezebben megy a nyelvtanulás. Még kérdéses, hogy mikor és milyen okoknál fogva lépnek fel a nyelvelsajátítást gátló tényezők (sorompók). A szakiroda-

mélet is magyarázza: (i) a **biológiai sorompó** elmélete a csökkenő agykérgi plaszticitást okolja a felmerülő nehézségekért; (ii) a **kognitív sorompó** elmélete szerint a tizenkét éves kor körül kialakuló elvont gondolkodás gátolja az „öntudatlan”, gyermekkori nyelvtanulást; és (iii) az **affektív sorompó** elmélete, amely szerint a pubertás korban végbemenő személyiségváltozások következtében a gyermek már nehezebben fogadja el az idegen nyelvi normát (vö. Kassai 1995, 109). A három teória közül a biológiai sorompóval foglalkozó tűnik a legelfogadhatóbbnak, bár ennél számos vita folyik még arról, hogy a nyelvtanulást gátló második biológiai sorompó mikor is következik be pontosan. Az mindenesetre tényként fogadható el, bár az objektív hatások mellett számos szubjektív tényező is befolyásolhatja a nyelvtanulásmenetét és eredményességét, hogy a kritikus periódus eltelte után hol kisebb, hol nagyobb mértékben, de megjelenik az akcentus. Az idegen nyelv hangtanának elsajátítása folyamán számos gond jelentkezik, mind az artikuláció, mind a percepció terén. Ilyenek például bizonyos hangszegmentumok és hangkapcsolatok hibás artikulációja, nem megfelelő hangsúlyozás és intonáció, ezeknek következtében az anyanyelvi hallgatók nehezen értik meg az akcentusos beszédet. Másrésztől a tanulók idegennyelv-percepciója is zavart – ők sem értik meg, amit az anyanyelvi beszélő mond (Szimeonova 1979, 16).

Az akcentusos beszéd kutatásának, az elméleti vizsgálatok mellett, számos gyakorlati eredménye is fellelhető. Barry a német anyanyelvű diákok angol kiejtését elemezte, párhuzamosan végzett a beszédprodukción és a percepción ellenőrző kísérleteket is (1989). Megállapításai szerint nemcsak általános törvényszerűségek mutathatók ki, de a különböző résztvevők teljesítménye között is nagy eltérések jelentkeznek.

Sopova és Veleva bolgár kutatók spanyol anyanyelvű egyetemi hallgatók bolgár beszédének magánhangzóit vizsgálták. A legtöbb gondot azok a hangok és hangtani jelenségek okozzák, amelyek nem lelhetők fel az anyanyelv artikulációs bázisában (1989).

A magyar anyanyelv okozta akcentussal (pontosabban annak intonációra gyakorolt hatásával) foglalkozik Pető (1980) tanulmánya, aki magyar egyetemi hallgatók orosz beszédének hanglejtési sajátosságait elemezte.



A jelen dolgozat a szláv anyanyelvű személyek magyar beszédének fonetikai jellegzetességeit vizsgálja.

### **Anyag és módszer**

Az akcentus jelenségének elemzése céljából két egymással összefüggő kísérletet végeztünk.

I. Az első kísérletben három szláv anyanyelvű személy (bolgár, orosz, lengyel) vett részt. Mind a három résztvevő már évtizedek óta Magyarországon lakik, és rendszeresen használja a magyar nyelvet munkájában és a mindennapi kommunikációban. Közös jellegzetességük az, hogy viszonylag későn (25 éves kor után) kezdték el a magyar nyelv elsajátítását, így beszédükben jól felismerhető az anyanyelv(ek) hatása. A személyek kiválasztásánál az is szempont volt, hogy mind a három szláv csoport (déli, keleti és nyugati) képviselve legyen, hiszen így jól megfigyelhetők a közös és az eltérő hangtani jelenségek. A résztvevők átlagéletkora 55 év, és mind a hárman férfiak. Interjúkat készítettünk velük (spontán beszéd), 15-20 perces időtartamban, amelyeket magnetofonra rögzítettünk. A felvett anyagokat többször lehallgattuk, kiválasztottuk az elemzésre kerülő egységeket. Miután a szláv nyelvek mássalhangzó-rendszere nem különbözik lényegesen a magyar nyelvétől, feltételeztük, hogy az akcentus kialakulásáért elsősorban a magánhangzó-rendszerek közötti eltérések okolhatók. Ezért öt olyan viszonylag gyakori magyar magánhangzót választottunk ki, amelyek nem, vagy csak részben találhatók meg a három vizsgált szláv nyelvben, így viszonylag jól megfigyelhető az anyanyelvi artikulációs bázis hatása. Ezek az [ɔ], [a:], [e], [e:] és [ø] magánhangzók voltak. A kiválasztott magánhangzókból a három beszélőnél 15-15-öt elemeztünk, összesen mintegy 600 adatot dolgoztunk fel.

Az akusztikai-fonetikai vizsgálat tárgya a magánhangzók formánsstruktúrájának eltérései a normatív magyar kiejtéshez viszonyítva. Miután egy adott hang legjellegzetesebb tulajdonságait az első és a második formáns határozza meg, ezért ezeket vizsgáltuk meg CSL-4300B digitális jelelemző készülék segítségével. A kapott adatokat átlagoltuk, majd összevettük mind a magyar magánhangzók, mind pedig a három szláv nyelv magánhangzóinak átlagos formánsértékeivel.



II. A második kísérletben percepció vizsgálatot végeztünk 10 magyar anyanyelvű kísérleti személlyel. A résztvevők között 6 nő, 4 férfi volt, átlagéletkoruk 35 év. A meglévő interjúkból 20-20 másodperces részleteket vágtunk ki, amelyeket lejátszottunk a magyar anyanyelvű személyeknek. Arra kértük őket, hogy kétszeri meghallgatás után mondják meg, feltételezésük szerint, vajon mi lehet a beszélők anyanyelve. Arra is megkértük őket, hogy 1-től 5-ig osztályozzák a beszédprodukciók minőségét. A kísérlet célja, hogy adatokat kapjunk arról, mennyire lehet az akcentus alapján megállapítani a beszélő anyanyelvét, illetve, hogy megtudjuk, mi zavarja jobban a hallgatót – a beszéd fonetikai hibái vagy a nyelvtani pontatlanságok.

### **Kísérleti eredmények**

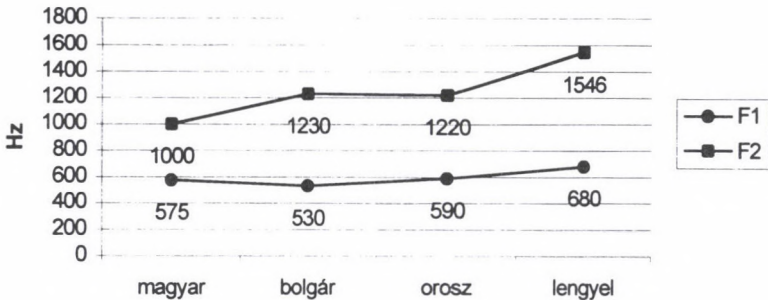
Az első kísérlet eredményeit grafikonok és táblázatok segítségével foglaltuk össze, továbbá összevetettük mind a bolgár, az orosz és a lengyel, mind pedig a magyar magánhangzók formánsainak átlagértékeivel.

A három szláv nyelv magánhangzó-rendszerének vannak közös, illetve egymástól erősen eltérő jellemzői is. Közös az alap-magánhangzók [a, e, i, o, u], illetve az a tény, hogy hiányzik a magyar nyelvben előforduló számos magánhangzó [ɔ, e:, o:, ø, ø:, i:, u:, y:]. Az eltérések között ki lehet emelni a hangsúlyos–hangsúlytalan ellentétet, amely kimutatható mind a bolgár, mind pedig az orosz nyelvben, azonban más-más eredménnyel jár a magánhangzók minőségére és az értelem-megkülönböztető szerepre nézve. A lengyel magánhangzó-rendszert a nazális magánhangzók megléte különbözteti meg a másik két vizsgált nyelvtől (Bańcerowski 1980, 127). Ennek ellenére a magánhangzó-formánsok átlagértékei nem különböznek lényegesen egymástól (1. táblázat).

1. táblázat: A vizsgált szláv nyelvek magánhangzóinak formáns-  
struktúrája (az oroszra – Bolla 1981; lengyelre – Bolla-Földi 1987;  
bolgárra – Sopova-Veleva 1989)

	Bolgár		Orosz		Lengyel	
Hang	F1	F2	F1	F2	F1	F2
<b>a</b>	743	1210	747	1470	650	1200
<b>e</b>	491	2192	400	1957	570	2075
<b>i</b>	231	2477	309	2045	230	1950
<b>o</b>	499	945	419	1026	540	1200
<b>u</b>	262	687	309	875	335	850
<b>ə (ɐ)</b>	446	1423	460	1450	–	–

Az [ɔ] hang esetében igen jelentős különbségekre lehet számítani, hiszen mind a három vizsgált szláv nyelvben csak egyfajta *a*-hang található, amely lényegesen eltér a magyar [ɔ] és [a:] hangoktól. A mérések eredményeit az 1. ábra mutatja.



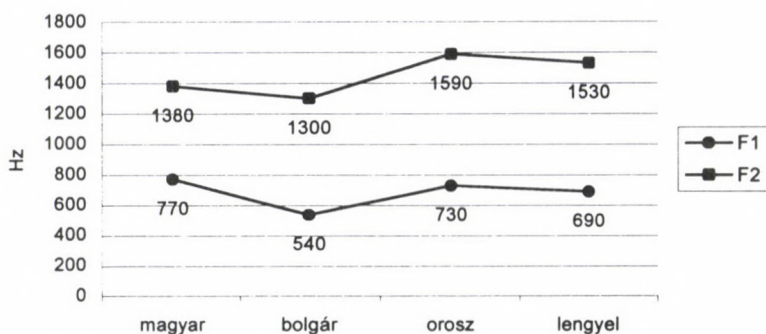
1. ábra

Az [ɔ] hang formánsai magyar és szláv kiejtésben

A 1. ábrán látható adatok jól mutatják a magyar és a szláv nyelvek artikulációs bázisainak eltéréseit. Különbségek mutathatók ki az [ɔ]

hang kiejtésében mind az első, mind a második formáns tekintetében. Legjelentősebb a lengyel beszélőnél jelentkező különbség, amely az F1-nél több mint 100 Hz, az F2-nél viszont már 546 Hz. A bolgár és az orosz beszélő [ɔ] hangjai átlagban nem térnek el jelentősen egymástól, a magyar [ɔ] hanghoz viszonyítva csak az F2-nél tapasztalható jelentősebb (200 Hz feletti) különbség. (Az ábrákon látható magyar magánhangzó-formánsok forrása Bolla 1978.)

Hasonló eredményeket kaptunk az [a:] hang elemzése során is (2. ábra).



2. ábra

Az [a:] hang formánsai magyar és szláv kiejtésben

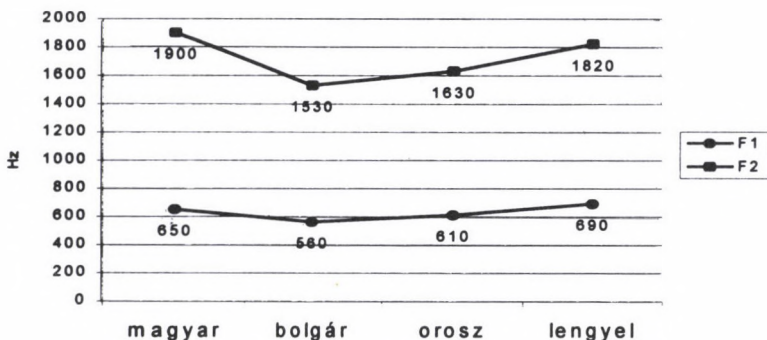
Az [a:] hang esetében még jelentősebb különbségeket tapasztalhatunk, hiszen mind három beszélőnél, vagy az első vagy a második formánsnál 200 Hz körüli eltérés mutatkozik.

Ha összevetjük az [ɔ] és az [a:] hangok formánsátlagait a vizsgált szláv nyelvek normatív kiejtésének átlagértékeivel, szintén megfigyelhetünk bizonyos eltéréseket. A bolgár beszélőnél mind az [ɔ], mind az [a:] hang első formánsa jelentősen (200 Hz körül) eltér a bolgár [a] átlagától (vö. 1. táblázat), és inkább a bolgár [o] hanghoz esik közelebb. A második formáns esetében ez a különbség már nem



annyira szembetűnő. Az orosz anyanyelvű [ɔ] hangjának átlagértéke erősen eltér az orosz [a] hangétól, viszont ez az [a:] -ra már nem érvényes. A lengyel beszélő jellegzetessége, hogy az általa ejtett [ɔ] és [a:] hangok formánsértékei között alig tapasztalható eltérés: [ɔ] F1=680, F2=1546; [a:] F1=690, F2=1530, tehát nála egyfajta „átlag” [a] hang hallható, amely viszont erősen eltér (főleg a második formánsban) a lengyel normatív kiejtés átlagos formánsértékétől.

Miután mindhárom vizsgált nyelvben megtalálható az [e] hang, azt vártuk, hogy a kísérletben részt vevő személyeknek ez fogja a legkisebb gondot okozni, vagyis az [e] hang kiejtésénél fogjuk a legkisebb eltéréseket tapasztalni (3. ábra).

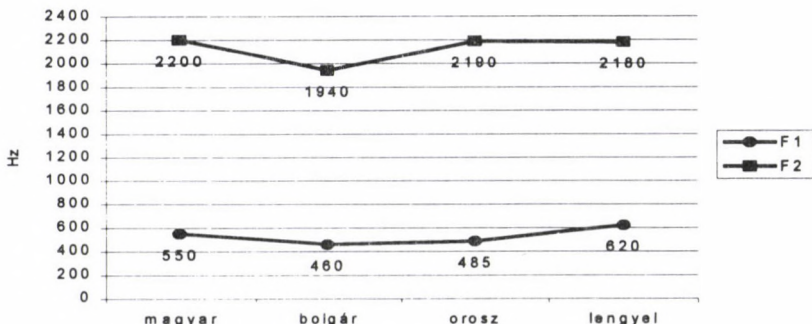


3. ábra

Az [e] hang formánsai magyar és szláv kiejtésben

Hipotézisünk csak részben igazolódott be. Az [e] hang első formánsánál valóban kisebb eltéréseket tapasztaltunk a magyar kiejtéshez viszonyítva. A második formánsnál azonban, különösen a bolgár és az orosz beszélőnél a különbség már jelentős (a bolgárnál 370, az orosznál 270 Hz). A lengyel beszélő [e] hangja szinte teljesen megegyezik a magyar átlaggal.

Az [e:] hangnál azt vártuk, hogy a lengyel és főleg az orosz beszélőnél csekély eltérések jönnek létre, míg a bolgár beszélőnél, ahol az anyanyelvben a magánhangzók palatalizációja gyengébb (Lilov 1972), gondot okozhat a helyes kiejtés. A kapott eredményeket a 4. ábra foglalja össze.



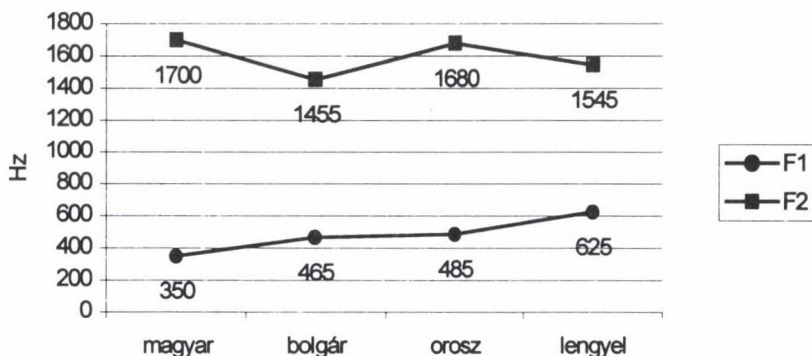
4. ábra

Az [e:] hang formánsai magyar és szláv kiejtésben

Az eredmények ezúttal alátámasztották feltételezésünket – mind az orosz, mind a lengyel beszélő szinte tökéletesen ejtette a magyar [e:] hangot. A bolgár résztvevőnél tapasztalható volt bizonyos eltérés, amely az első formáns estében nem nagy, a második formánsnál viszont már jelentősebb volt (260 Hz).

Ha összevetjük a három szláv nyelv [e] hangjait, azonnal feltűnik, hogy alig mutatható ki közöttük különbség (vö. 1. táblázat). Ezek az értékek azonban sokkal közelebb állnak a magyar [e:] hanghoz, mint az [e]-hez, holott szubjektíve (hallás után) azt gondolnánk, hogy ennek pont a fordítottja igaz. Azzal azonban mindenképpen számolni kell, hogy a szakirodalomban közölt értékek az átlagot képviselik, így nem mindig lehet segítségükkel az árnyalatnyi különbségeket kifejezni.

Feltételezésünk szerint az utolsó vizsgált hangnál lehetett a legnagyobb eltérésekre számítani, hiszen az [ø] hang egyik vizsgált szláv nyelvben sem található meg. Az eredményeket az 5. ábra mutatja.



5. ábra

Az [ø] hang formánsai magyar és szláv kiejtésben

A kapott eredmények itt is igazolták a kiinduló feltételezést. Ha egy hang hiányzik az anyanyelvi artikulációs bázisból, viszonylag nehéz azt elsajátítani, főleg felnőtt korban. Az adatok érdekessége, hogy mind a három beszélő különböző módon hozza létre az [ø] hangot, de egyikőjük sem képes helyesen vagy a helyeshez közel létrehozni. Az a tény azonban, hogy a formánsok tekintetében a kiejtett hangok nem egyeznek meg a magyar nyelvben lévő [ø] hangéval, alapvetően nem okoz megértési problémákat – a hang a hallgató számára mindig egyértelműen azonosítható.

Adataink igazolták, hogy a nem anyanyelvi beszélők számára leginkább azok a beszédhangok okoznak gondot, amelyek hiányoznak az anyanyelvi artikulációs bázisból. Emellett a beszélőnek figyelnie kell a magyar hangok időtartamára is, de gyakran még azokat a hangokat sem tudja pontosan kiejteni, amelyek elvileg előfordulnak anyanyelvében.



sem tudja pontosan kiejteni, amelyek elvileg előfordulnak anyanyelvében.

Miután a közölt adatok átlagértékek, nem tükrözik teljesen a három résztvevő által kiejtett magánhangzók formánsainak határértékeit, amelyeknek segítségével sokkal árnyaltabb képet kaphatunk az akcentus lényegéről.

2. táblázat: A formánsok határértékei a szláv anyanyelvű beszélők kiejtésében

	<b>Bolgár</b>		<b>Orosz</b>		<b>Lengyel</b>	
Hang	F1	F2	F1	F2	F1	F2
[ɔ]	483-604	892-1606	520-637	989-1511	610-859	1237-1704
[a:]	465-621	1047-1811	585-807	1269-1914	627-859	1368-1770
[e]	449-639	1345-1746	481-690	1189-2186	514-885	1623-2057
[e:]	336-550	1746-2018	390-578	1913-2427	498-807	2005-2473
[ø]	413-639	1358-1642	337-585	1210-1874	498-807	1237-1875

A 2. táblázat adataiból jól látszik, hogy a vizsgált magánhangzók igen tág értékhatárok között mozognak mind a három beszélőnél (az első formánsnál 150-200 Hz között, a másodiknál azonban akár 600 Hz eltérést is találunk ugyanazon hang különböző megvalósulásainál). Ez azt jelenti, hogy a beszélők kiejtése bizonytalan, gyakran nem találják el pontosan a hangot, illetve akaratukon kívül más ejtenek helyette. Ez a jelenség is hozzájárul ahhoz, hogy az anyanyelvi hallgató akcentusosnak érzékelje beszédüket, sőt néhány esetben a folyton változó kiejtés a megértést is zavarhatja. A leírtak jól mutatják az anyanyelvi artikulációs és percepciós bázis akadályozó hatását a felnőttkori idegennyelv-elsajátításkor. Mindhárom résztvevő (saját elmondásuk szerint) auditív úton, utánzással sajátította el a magyar nyelvet. A már meglévő percepciós bázisuk zavarta a magyar nyelv jellemző hangjainak észlelését és elsajátítását, míg az anyanyelvi artikulációs bázis a már megtanult hangok helyes kiejtését tette gyakorlatilag lehetetlenné. Ennek eredményeként keletkezett a beszélők jellegzetes akcentusa.

hallgatás után a résztvevőknek el kellett mondaniuk feltételezéseiket a beszélők anyanyelvéről, a kapott adatok igen változatosak lettek. Említették az angol, az arab, a francia és a német nyelvet, illetve különböző szláv nyelveket is (orosz, szlovák és „szláv”). A szláv nyelvek aránya az első meghallgatás után meghaladta az 50%-ot, tehát a magyar anyanyelvű kísérleti személyek számára viszonylag jól felismerhető a szláv anyanyelvűek jellegzetes akcentusa. A felvételek második meghallgatása előtt a kísérletvezető közölte, hogy mindhárom beszélő szláv anyanyelvű. Az eredményeket a 3. táblázatban foglaltuk össze, továbbá itt közöljük a résztvevőknek a beszédprodukciókra adott osztályzatait is.

3. táblázat: A beszélők anyanyelvének felismerési aránya a második meghallgatás után

<b>Bolgár beszélő</b>	<b>Orosz beszélő</b>	<b>Lengyel beszélő</b>
orosz – 8 fő	lengyel – 3 fő	szlovák – 6 fő
ukrán – 1 fő	cseh – 2 fő	cseh – 1 fő
<b>bolgár – 1 fő</b>	<b>orosz – 2 fő</b>	<b>lengyel – 1 fő</b>
	szlovák – 1 fő	
	horvát – 1 fő	
osztályzat: 3, 7	osztályzat: 3, 4	osztályzat: 3, 2

A kapott adatok között feltűnően kevés a jó eredmény. Ez a bolgár beszélő esetén egyáltalán nem meglepő, hiszen Magyarországon viszonylag kevés lehetőség adódik arra, hogy ilyen típusú akcentussal találkozzunk. Az „ukrán” megjelölése, mint potenciális anyanyelv leginkább a meghatározhatatlanságot jelenti, hiszen ukrán nyelvet gyakorlatilag alig hallani. Igen érdekes az a tény, hogy a tíz hallgatóból nyolc orosznak gondolta a bolgár személyt, arra viszont, aki valóban az orosz volt, csak ketten „tippeltek” helyesen. Az orosz beszélőnél figyelhetjük meg a legváltozatosabb válaszokat. Nála gyakorlatilag a Magyarországon közsímbert összes szláv nyelv megjelenik, és a hallgatók nagyobbik része a nyugati szláv csoporthoz tartozónak ítéli őt. A lengyel személy anyanyelvi hovatartozására adott válaszokat tekint-



hetjük a legpontosabbnak, hiszen mindegyik hallgató a nyugati szláv csoportba helyezi őt beszéde alapján. Nála is azonban előtérbe kerülnek azok a szláv nyelvek (szlovák és cseh), amelyekről a magyar hallgatónak lehet valamilyen benyomása.

A beszédprodukciónak osztályozásánál feltűnő volt, hogy az osztályzatok döntő többsége 3-as és 4-es volt, csak egy esetben adtak 2-est, míg 5-ös egyáltalán nem fordult elő. Ez végül is azt jelenti, hogy mind a három beszélő produkcióját közepesnek vagy jónak ítélték, tehát az akcentus mértéke még elfogadható volt a magyar hallgatók számára. A lejátszott részletekben elenyésző a nyelvtani hibák aránya, a beszélők beszédtempója viszont eltérő, így az akcentus mellett ez is befolyásolta a döntést. A bolgár személy beszélt a leglassabban, ügyelve arra, hogy ne kövessen el nyelvtani hibát (ami nem mindig sikerült), és ezzel magyarázható az a tény, hogy ő kapta a legjobb osztályzatot. Az orosz és a lengyel résztvevő minősítése között gyakorlatilag alig van különbség, mind a ketten többségében közepes osztályzatokat kaptak. Valószínűsíthető, hogy azok a hallgatók, aki közepes osztályzatot adtak inkább a beszédprodukcióna (fonetikai jellemzők) koncentráltak, míg azok, akik jó jegyet adtak, inkább az érthetőséget és a nyelvtani jártasságot helyezték előtérbe.

### Befejezés

Az elvégzett kísérletek segítségével egyértelműen bizonyítást nyert az a feltételezés, hogy az idegen anyanyelvűek beszédének fonetikai sajátosságai műszeres úton mérhetőek, ami által lehetőség nyílik arra, hogy a szubjektíven (hallás útján) szerzett benyomásokat objektíven (mérési adatokkal) is alá tudjuk támasztani. A kapott adatok elemzésével az is világossá vált, a magyar nyelv számos jellegzetes hangtani sajátosságai közül melyek okolhatók leginkább a kísérletben részt vevő szláv anyanyelvűek beszédének akcentusáért. Az akcentus-élmény létrejöttéért elsősorban azok a magánhangzók felelnek, amelyek nem lelhetők fel a kísérleti személyek anyanyelvének artikulációs bázisában. Azok az objektív adatok (formánsátlagok), amelyeket a kísérlet folyamán kaptunk, nem mindig tértek el lényegesen a magyar átlagoktól, de ha a mért értékek szórási arányát is figyelembe vesszük, akkor már jelentősek a kimutatható különbségek. Összefoglalásként



elmondható, hogy a szláv anyanyelvű beszélők magyar beszédének akcentusáért egyszerre felelősek a formánsok eltérései és a nem állandó artikuláció. Ezek mellett természetesen léteznek más tényezők is (intonáció, a magánhangzók időtartama stb.), amelyek szintén szerepet játszanak az akcentus kialakulásában.

A percepciós kísérlet segítségével arról is kaptunk adatokat, hogyan élük meg és ítélik meg az anyanyelvi hallgatók az akcentus megletétét, illetve mennyire képesek következtetni az eredeti anyanyelvre az akcentus jellege alapján. Bár a kapott eredményekben feltűnően kevés volt a jó válasz, a résztvevőknek mégis sikerült legalább viszonylagos pontossággal (nyelvcsalád, földrajzi terület) megállapítani a beszélők anyanyelvét. A beszédprodukció megítélésénél valószínűleg két tényező – a kiejtés pontossága és az érthetőség játszotta a legfontosabb szerepet. Ezzel magyarázható az a tény, hogy a hallgatók gyakorlatilag csak kétfajta osztályzatot (közepes és jó) adtak.

Az akcentusos beszéd további kutatása még számos adattal segítheti az idegennyelv-oktatást és azoknak a külföldieknek a jobb beilleszkedését a befogadó társadalomba, akik valamilyen okból kifolyólag huzamosabb időt töltenek ott.

### Irodalom

Bañcerowski Janusz: A lengyel és a magyar vokális rendszer egybevetésének néhány kérdése. Magyar Fonetikai Füzetek 5. 1980, 117-128.

Barry, W. J.: Perception and production of english vowels by German learners. *Phonetica* 46/4. 1989, 155-168.

Bolla Kálmán: A magyar magánhangzók analízise és szintézise. Magyar Fonetikai Füzetek 1. 1978, 53-68.

Bolla, K.: A conspectus of russian speech sounds. Akadémiai Kiadó. Budapest 1981.

Bolla, K.–Földi, É.: A phonetic conspectus of polish. Magyar Fonetikai Füzetek 18. 1987.

Gósy Mária: Beszédszlelés. MTA Nyelvtudományi Intézete. Budapest 1989.

Kassai Ilona: Pszicho-szociolingvisztikai jegyzetek az akcentusról. ÁNyT 18. 1995, 103-115.

Lilov, M.: Foneticsnite osnovi na govornata posztanovka v balgarszkija ezik. *Balgarszki ezik* 3. 1972, 210-220.

Pető Zsigmond: Intonációs típushibák a főiskolai hallgatók orosz beszédében. Magyar Fonetikai Füzetek 5. 1989, 164-168.

Sopova, J.–Veleva, M.: Fonetiko-fonologicsni greski pri vazpriemaneto na balgarszkata vokalna szisztema ot iszpanogovoresti sztudenti. Szaposztavitelno ezikoznanie 1. 1989, 17-24.

Szimeonova, R.: Za definicijata i szadarzsanieto na ponjatieto „artikulacionna baza” na daden ezik. Szaposztavitelno ezikoznanie 1. 1979, 15-20.

# MAGÁNHANGZÓK KONTRASZTÍV AKUSZTIKAI ELEMZÉSE

**Brenner Koloman**

**Veszprémi Egyetem, Német Nyelv és Irodalom Tanszék**

## **Bevezetés**

A beszéd hangalakjának akusztikai elemzése fontos adatokat szolgáltat azon paraméterekről, amelyek a percepció első, akusztikai szintjén meghatározzák a beszédjel feldolgozásának módját az adott személynél. Ezen a szinten a hangalak felbontásra kerül, mégpedig akusztikai részeire: például az időviszonyok, a frekvencia jellemzői, vagy az intenzitás (Valaczkai 1989, 389). A percepció ezen első szintjének létezését több kísérlet is bizonyítja, így például Pisoni négy bilabiális hangot játszott le mássalhangzó–magánhangzó kapcsolatban a kísérleti személyeknek (a magánhangzó a rövid [a] típus volt), amelyek különbözőségét kizárólag a 20 ms kiindulási nagyságról mindig 20 ms-mal megnövelt időtartam adta, amely után a zöngésség szakasza bekövetkezett. A résztvevők feladata a mássalhangzó-típusok azonosítása volt. A kísérlet eredményeképpen a kísérleti személyek meg tudták ugyan a négy mássalhangzót különböztetni, de megnevezni nem sikerült mindegyiket (Gósy 1981, 89).

A jelen tanulmány célja, hogy az [a:], [u:] és [i] hangtípusok fontos akusztikai paramétereit konfrontatív–kontrasztív módszerrel egybevesse, egy nyugat-magyarországi német dialektus, a normatív német, illetve a magyar kiejtés tekintetében. A beszédjelből szegmentált hangtípusok akusztikai összetevőinek vizsgálata elősegíti, hogy az adott szegmentumok tipizálható adatait összehasonlítva, egy első, alapvető áttekintést adjunk a releváns megfelelésekről, illetve különbözőségekről. Az [a:], [u:] és [i] hangtípusok mindhárom vizsgált esetben a magánhangzó-rendszer egyértelműen azonosítható és meghatározó elemei közé tartoznak. Izolált szavakban megjelenő



hangtípusok akusztikai paramétereivel dolgozunk, mivel a metodológiai sajátosságok miatt csak ezekről rendelkezünk egybevethető, megbízható adatokkal. A magánhangzók minőségének megváltozása a spontán beszédben csak a vizsgálatok egy következő lépésében kerülhet feldolgozásra (vö.: Gósy 1996). A vizsgálat tárgya tehát nyugat-magyarországi keletdunai-bajor (ostdonaubairisch) dialektust kompetensen beszélők kiejtése, összevetve a normatív német, valamint a magyar kiejtéssel. A normatív német és magyar magánhangzók adatai Valaczkai (1998) és Szalai (1995) vizsgálataiból származnak.

Ezen felül a tanulmány célja, hogy kiderítse: észlelhető-e a nyugat-magyarországi német dialektust beszélő személy kiejtésében a magyar vonatkozásában interferencia, illetve, hogy – főleg a műholdas német nyelvű elektronikus médiumok elterjedése révén – az utóbbi évtizedben erősödő normatív német kiejtés behatása mennyire nyomon követhető. Ezen kívül az általános megfelelési és a specifikálható adatok egybevetése is megtörténik.

### **A nyugat-magyarországi német dialektusok**

A magyarországi német dialektusok között az osztrák határmentén használt keletdunai-bajor, avagy keletközép-bajor nyelvjárásformák különleges szerepet töltenek be. Ezt egyrészt a földrajzi helyzet indokolja, itt gyakorlatilag a zárt német nyelvterület közvetlen szomszédságában élnek a beszélők, s még az elmúlt évtizedek elzárt-sága sem tudott bizonyos rokoni, médiabeli kapcsolatokat megszüntetni (a felmérések szerint a nyugat-magyarországi németek körében már a korábbi évtizedekben is az egyszerűen fogható osztrák televízió, az ORF adásai voltak a legnépszerűbbek; vö.: Manherz 1977, 128). Másrészt az itt élő német kisebbséghez tartozók ősei nem a török dúlás után kerültek Magyarországra, hanem gyakorlatilag a XIII–XIV. századtól az ország ezen részén már néhol többségbe kerültek. Így sokkal régebbi gyökerekkel rendelkeznek, mint a délkelet-dunántúli németiség, amely a mai Magyarországon élő német közösség zömét alkotja. E történeti és szociális háttér befolyásolta és befolyásolja

a nyugat-magyarországi német ún. „települési dialektusok“ jellemzőit is.

Települési dialektusnak nevezzük azokat a nyelvjárásokat, amelyek csak a betelepült németek új hazájában alakultak ki véglegesen, a különböző nyelvjárásokat beszélő lakosság következtében rövidebb-hosszabb ideig tartó nyelvi kiegyenlítődési szakaszok után. E nyelvjárások keveredése és kiegyenlítődése után Nyugat-Magyarországon annyiban sajátos helyzet alakult ki, hogy az ún. „rendező“ nyelvjárásforma – amely itt keletdunai-bajor ui-dialektust beszélők révén jött létre – után a nagyobb régió rendező nyelvjárása, a bécsi köznyelv is érvényesítette befolyását (Hutterer 1959–1961, 53).

A vizsgálat a közigazgatásilag Sopronhoz tartozó Brennberg településről származó adatokat használja fel (Brenner 1994), amelyek a „Voice Identification Incorporation“ (Series 700) készülék segítségével készült spektrális elemzésből származnak. Az itt beszélt keletdunai-bajor dialektus jellemző a nyugat-magyarországi régió egészére nézve is. Mivel az 1980-as évek végéig a település a határsávba tartozott, és így meglehetősen elszigetelt volt, a németül beszélők nyelvi váltása a magyar felé nem zajlott le olyan mértékben, mint a legtöbb magyarországi településen.

A német dialektus legfontosabb jellemzői közül a következőket kell kiemelni: a Sopron környéki német dialektusokra általánosan jellemzően, az ófelnémet-óbajor (althochdeutsch-altbairisch) ai diftongus oa formában jelenik meg, például [oÁ] „Ei“, ezen kívül az erős diftongizálást kell megemlíteni, különösen nazális és likvida hangtípusok előtti pozícióban. A régi [l] hang előtt a magánhangzó-típusoknál fellelhető az ajakkerekítési tendencia, ez egyébiránt csaknem az összes ausztriai német dialektusra is jellemző, például [hy:f] „Hilfe“. Az egész bajor nyelvjárás területre jellemző tendencia az l és n hangok palatalizálása, ez is fellelhető a kérdéses dialektusban: [fʁɪd] „Feld“.

A német dialektusok felosztása szempontjából döntő a második mássalhangzó-eltolódás jelenségeinek megléte, illetve az egyes változások részleges, avagy teljes végrehajtása. Mint a felsőnémet dialektusoknál mindig, a nyugat-magyarországi német nyelvjárások is a má-



sodik mássalhangzó-eltolódás teljes és konzekvens végrehajtását mutatják: például [dik<sup>a</sup>opf] „Dickkopf“ (germ. p > ófn. pf); [tsaunt] „Zahn“ (germ. t > ófn. (t)z) (Brenner 1994, 10).

### **A normatív német kiejtés**

Az írását tekintve nagyjából egységes német nyelv körülbelül a XVII. századra alakul ki, egy rendkívül hosszú és komplex fejlődési folyamat során. Hosszú évszázadokon keresztül a német nyelv egyetlen beszélt formájának tekinthető dialektusok szabták meg az írásos formát is. Már az ófelnémet kortól kezdődően érezhető természetesen az egységesülési tendencia is, amely azonban nem volt lineáris fejlődését tekintve, a XIII–XIV. században például ismét felerősödött a dialektusok befolyása az írásképre. A különböző politikai, társadalmi és kulturális változások az egységesülés felé vitték az írásos formát, a német nyelv azonban a mai napig őrzi pluricentrikus jellegét, és ez különösen igaz a kiejtésre.

Alapvető megállapítás, hogy a „tisztá“ normatív kiejtés a köznapi életben csak elvétve fordul elő. Tekintsük át röviden a normatív német kiejtés felé vezető utat! A német klasszicizmus időszakában Goethe és Schiller tette az első lépéseket a színházi beszéd egységesítése érdekében. Az egységesítés első tudományosan kidolgozott alapelvét Hermann Paul említi először „Prinzipien der Sprachgeschichte“ (1909, első kiadás 1880) című művében. E szerint nincs olyan földrajzi terület a német nyelvterületen belül, amely az egységesítés alapjaként szolgálhatna, ezért a komoly drámai színészek kiejtését kellene figyelembe venni. Ezt az alapelvet figyelembe véve hajtja végre Theodor Siebs (1969) és bizottsága az első kísérletet az egységes kiejtés rögzítésére. Az ún. „színpadi kiejtés“ (Bühnenaussprache) alapján állítják össze 1898-ban az első munkát, amely szöszedetet is tartalmaz. Az egyes szóalakok kiejtését általában a felnémet nyelvi alakokat alnémet hangértékekkel ellátó alapelv alapján határozták meg. Ez a norma azonban soha nem tudott a német nyelvi közösség széles köreiben elterjedni, mivel egy szinte abszolutizált, elérhetetlen-



nek tűnő normát határozott meg, amelyet manapság irreálisnak, túlzónak, patetikusnak érzünk.

A második és egyben jelenleg is érvényes kodifikáció alapjául a rádió- és tévébemondók által használt ún. „Rundfunkaussprache” szolgált. A készülékek gyors elterjedése után a német nyelvi közösség valamennyi rétegéhez eljut ez a kiejtésfajta, és – a siebs-i normával ellentétben – akkor is elfogadják normának, ha maguk a kompetens német beszélők nem is tudják kiejteni. E normát először a „Wörterbuch der deutschen Aussprache” (első kiadás 1964, Lipcse) szerzőkollektívája kodifikálja, miután empirikus módszerrel feltérképezik a hivatásos bemondók kiejtésének jellemzőit. A legismertebb német nyelvvel kapcsolatos átfogó jellegű mű, a Duden (1974) megfelelő kötetében 1962-ben még a siebs-i formák szerepelnek, s a későbbiekben is ragaszkodnak a színházi kiejtés elsőrangú fontosságához. Ennek ellenére az 1974-es második kiadású Duden kiejtési szószerdetében már átveszik a „Wörterbuch”-ban tíz évvel korábban kodifikált formákat (Valaczkai 1998, 21-24).

Megemlíjtük, hogy a német nyelvi közösség jelentős része e normát nem tudja kiejteni (különösen az idősebb korosztályok azokon a területeken, ahol a dialektusok hatása erős), és a köznap életben vagy a nyelvjárást, vagy az ún. „közbeszéd”-et (Umgangssprache) – amely vegyíti a nyelvjárásokat és az irodalmi németet – beszélik a legtöbben.

A Valaczkai (1998) által elvégzett kísérletek eredményeit a következő körülmények befolyásolták: az adatközlő kiejtésében a Gräfen-thali német tájszólás nyomai ugyan fellelhetők voltak, de kiejtésének alapjául a fent ismertetett, a hivatásos bemondók által is használt norma szolgált. A vizsgálatot szintén a „Voice Identifikation Incorporation” (Series 700) készülékével végezték el.

### **A magyar hangtípusok adatai**

A magyar nyelv vonatkozásában is sor került a fontosabb akusztikai paraméterek megállapítására, például Bolla (1980) hangalbumában. Olaszy (1989) vizsgálata beszédszintézis útján ellenőrzött adato-

kat szolgáltatott. Jelen egybevetés Szalai (1995) adatain alapszik, mivel az érintett három kardinális magánhangzótípusokra koncentrált, azzal a megjegyzéssel, hogy az illabiális [a:] hang adatait teszi közzé, míg a nyugat-magyarországi német, illetve a normatív német adatoknál nincs különbségtétel [a:] és [ɑ:] között. E vizsgálat a CSL 4300 jelfeldolgozóval készült spektrális elemzéseken alapult.

### Az egybevető vizsgálat eredményei

A nyugat-magyarországi német dialektusban, a normatív német, illetve magyar kiejtésben előforduló [ɑ:]/[a:], [u:] és [i] hangtípusokat reprezentáló szegmensek akusztikai paraméterei közül – mivel izolált szavakból szegmentálták őket – a percpiciós folyamatban meghatározó első három formáns határértékeinek egybevetését végeztük el:

1. táblázat: Az [ɑ:]/[a:] hangtípus(ok) adatai

<b>Brennberg</b> (Brenner)		<b>Normatív német</b> (Valaczkai)		<b>Magyar</b> (Szalai)	
F1	290-390 Hz	F1	610-850 Hz	F1	665-892 Hz
F2	730-760 Hz	F2	1180-1900 Hz	F2	1300-1496Hz
F3	1780-1930 Hz	F3	2980-3100 Hz	F3	2340-2971 Hz

Az [ɑ:] hangtípus adatainál a következő tendenciákat állapítottuk meg: a nyugat-magyarországi német tájszólás vonatkozásában feltűnő mindhárom formánsnál az abszolút értékek alacsonyabbak voltak, mind a normatív német, mind a magyar ejtéssel egybevetve. Ez részben az adatközlőtől függő faktorokra vezethető valószínűleg vissza, mivel a hangtípus alapfrekvenciája is alacsonyabb az átlagosnál: 70 Hz. Másrészt megállapítható, hogy a spektrális eloszlás hasonló képet mutat, a

formánsok értékeinek eltérése mellett. A normatív német és a brennbergi formánsértékek az F1 és F2 vonatkozásában több száz Hz-es eltérést mutatnak, az F3 több mint 1000 Hz-cel magasabb a normatív német hangtípusnál.

A normatív német, illetve magyar [ɑ:] / [a:] hangtípus adatai az F1 tekintetében szinte megegyeznek, az F2 esetében a normatív német felső értéke 1900 Hz, szemben a magyar 1496-os értékkel. Az F3-nál a normatív német értékek átlagosan 3000 Hz körül, míg a magyar alsó értékek 2340 Hz-nél találhatók meg.

2. táblázat: Az [u:] hangtípus adatai

<b>Brennberg</b> (Brenner)		<b>Normatív német</b> (Valaczkai)		<b>Magyar</b> (Szalai)	
F1	120-140 Hz	F1	160-400 Hz	F1	257-414 Hz
F2	350-470 Hz	F2	500-900 Hz	F2	637-803 Hz
F3	nincs adat	F3	nincs adat	F3	2228-2705 Hz

Az [u:] hangtípus adatainál kevésbé szembeszökő a nyugat-magyarországi hangtípus adatainak alacsonyabb volta, meg kell azonban említeni, hogy a magyar értékekkel egybevetve nagyobb a különbség, mint a normatív német vonatkozásában. A brennbergi felső értékek megint több száz Hz-cel elmaradnak a normatív német felső értékei mögött. A normatív német és magyar formánsértékek közelítő tendenciákat mutatnak, az F1-nél a normatív német alsó 160 Hz-es értéke, valamint az F2-nél az 500 Hz-es alsó értéke mutat nagyobb eltérést.

Az [i] hangtípus adatainál (3. táblázat) a következőket említjük meg: a nyugat-magyarországi német dialektus hangtípusa ismét preg-



nánsan alacsonyabb formánsértékeket mutat, a spektrális eloszlás itt nem igazodik olyan mértékben, mint az [a:] hangtípusnál.

3. táblázat: Az [i] hangtípus adatai

<b>Brennberg</b> (Brenner)		<b>Normatív német</b> (Valaczkai)		<b>Magyar</b> (Szalai)	
F1	130-170 Hz	F1	250-450 Hz	F1	252-382 Hz
F2	1360-1530 Hz	F2	2360-2560 Hz	F2	2179-2559 Hz
F3	1770-2200 Hz	F3	3100-3400 Hz	F3	2798-2951 Hz

Az F1 értéke 250-300 Hz-cel, az F2 értéke körülbelül 1000 Hz-cel, az F3 értéke 1300-1400 Hz-cel marad el átlagosan a normatív német hangtípus formánsainak értékeitől. A normatív német és magyar formánsértékek F1 és F2 vonatkozásában erőteljes konvergenciát mutatnak, egyedül az F3-nál mutatkozik nagyobb eltérés: a normatív német formáns 3400 Hz-es felső értéke lényegesen nagyobb, mint a magyar 2951 Hz-es érték.

Összefoglalásként megállapíthatjuk, hogy nyugat-magyarországi vizsgált hangtípusok adatai a magyar ejtés vonatkozásában szinte semmilyen megfelelést nem mutatnak. A normatív német vonatkozásában is csupán csekély megfelelési tendenciákat észlelhetünk. Ez arra utal, hogy nem észlelhetünk számottevő interferencia-jelenségeket a német nyelvjárást első nyelvként használóknál.

### Irodalom

Bolla Kálmán: Magyar Hangalbum. Magyar Fonetikai Füzetek 6. Budapest 1980.

Brenner, K.: Akustische Analyse der Brennbergischen deutschen Mundart. Szeged 1994. (Bölcsház doktori disszertáció)

Brenner, K.: Vergleichende Analyse der Plosive einer ungarndeutschen ostdonaubairischen Mundart und der normierten deutschen Hochlautung. *Studia Germanica Universitatis Vesprimiensis* 2/1. Veszprém 1998, 23-35.

Der Große Duden. Bd. 6. Aussprachewörterbuch. Wörterbuch der deutschen Standardaussprache. Mannheim, Wien, Zürich 1974.

Großes Wörterbuch der deutschen Aussprache. (Autorenkollektiv) Leipzig 1982.

Gósy Mária: A szegmentális hangszerkezet percepciójáról. Magyar Fonetikai Füzetek 8. Budapest 1981, 87-103.

Gósy Mária: A magánhangzók minősége a spontán beszéd szavaiban. In: *Beszéd kutatás '96*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1996, 1-19.

Hutterer, C. J.: Hochsprache und Mundart bei den Deutschen in Ungarn. Hochsprache und Mundart in Gebieten mit fremdsprachigen Bevölkerungsteilen. *Berichte über die Verhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig* 105/5. Berlin 1959-1961.

Manherz, K.: Sprachgeographie und Sprachsoziologie der deutschen Mundarten in Westungarn. Budapest 1977.

Olaszy Gábor: Elektronikus beszédelőállítás. Műszaki Könyvkiadó. Budapest 1989.

Paul, H.: *Prinzipien der Sprachgeschichte*. 4. Aufl. Halle 1909.

Siebs, Th.: *Deutsche Aussprache. Reine und gemäßigte Hochlautung mit Aussprachewörterbuch* 19. Berlin 1969.

Szalai Enikő: Az [u:], [a:] és [i] hangok koartikulációs mezőiről. In: *Beszéd kutatás '95*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1995, 83-92.

Valaczkai L.: *Atlas deutscher Sprachlaute*. Wien 1998.

# A BESZÉDRPRODUKCIÓ TUDATOS VÁLTOZTATÁSA: A BESZÉLŐ SZEMÉLY UTÁNZÁSA

Gósy Mária  
MTA Nyelvtudományi Intézete

## Bevezetés

Az alkalmazott fonetika egyik legújabb ága a beszélő személy egyértelmű, kétséget kizáró felismerésének akusztikai-fonetikai megoldási lehetőségeivel foglalkozik. Az utóbbi évtizedben jelentős eredmények születtek ezen a területen (Coulthard 1992; Schlichting–Sullivan 1998; Gósy–Nikléczy 1999). A megközelítések sokfélék, a matematikai számításoktól, az akusztikai méréseken át a szoros értelemben vett kísérleti-fonetikai és percepciós kísérletekig. Hagyományos terminussal a 'hangszínezet' meghatározása lenne a cél, pontosabban azoknak a paramétereknek a megjelölése, amelyek lehetővé teszik a beszélő azonosítását a beszéde alapján.

A jelen tanulmány egy sajátos helyzetben igyekszik a probléma megoldásához közelebb jutni. Ez a sajátos helyzet az, amikor valaki egy másik beszélőt személyesít meg azzal a határozott céllal, hogy a hallgatóban a másik személyt idézze fel. Általában csak tudatos beszélők, rendszerint színészek képesek ilyen fajta utánpótlásra (a legkülönbözőbb célokkal). Az utánpótló tudatosan vagy kevésbé tudatosan igyekszik olyan artikulációs mozgássorokat létrehozni, amelyek hangzásukban a másik beszélőre jellemző beszédet képviselik. Nem arról van tehát szó egyszerűen, hogy az utánpótló az utánpótló szándékozik személy beszédprodukciónak imitálni, ez a sok tekintetben eltérő artikulációs szervek és működtetésük következtében nem is lenne lehetséges. Az utánpótló az utánpótló beszéd hangzását igyekszik megvalósítani a saját artikulációs bázisán belül a saját artikulációs mozgásainak részleges módosításával. Az utánpótló helyzete azért is nehéz, mivel soha nem fogja úgy hallani az utánpótló beszéd akusztikumát, ahogyan azt a hallgató, hiszen a beszélő a csontvezetés révén is dekódol. Mik az utánpótló lehetőségei egy másik személy beszédének „reprodukálására”?



Az átlagos alaphangmagasság közelítése, az egyénien ejtett beszédhangok felismerése és artikulációja, a sajátos beszéddallam és hangsúlyozás, valamint a beszédtempó utánzása. Nézzük meg, hogy mindezek milyen feltételekkel valósíthatók meg (1. táblázat).

1. táblázat: A beszéd utánzásának feltételei (lehetőségek és korlátok)

Artikulációs tényezők	Az utánzás lehetőségei		
	fiziológiai tényező	működtetés	korlátozottság
alaphangmagasság	hangszalagok	változtatható	erős
rezonátor-rendszer	garat-, száj, orrüreg	változtatható	közepes
képzési konfiguráció	artikulációs szervek	változtatható	gyenge
beszéddallam	hangszalagok	változtatható	gyenge
hangsúlyozás	izmok	változtatható	alig
beszédtempó	artikulációs szervek	változtatható	alig

A táblázatban meghatározottak természetesen az ún. „tanult beszélőre” vonatkoznak, aki tudatosan képes az artikulációs mozgássorait követni és módosítani. Fontos ezt hangsúlyozni, hiszen az átlagos beszélőnek rendkívül nehéz feladatot jelent például a beszédsebesség változtatása; sokan egyáltalán nem is képesek rá. A leírtak azt szemléltetik, hogy az utánzó fiziológiai alkata, beszédszerveinek morfológiai meghatározottsága, valamint a saját beszédének begyakorlottsága különböző mértékben befolyásolják az utánzás sikerét. Nem véletlen, hogy az utánzások is különféleképpen hatásosak, az alig észrevehetőtől a feltűnően gyengéig. Mindez alapvetően függ a beszédet meghatározó objektív tényezőktől. A külső hasonlatosság (például arc, alkat) magában foglalhatja az artikulációs hasonlóságot is; ilyenkor az utánzónak könnyebb a dolga. Nem véletlen, hogy a színészek kiknek az utánzására vállalkoznak. Egyfelől olyan személyeket választanak, akiknek a beszédét a saját beszédprodukciójukkal jól meg tudják közelíteni; másfelől igyekeznek a sajátos hangzást eredményező és ismert beszélőket kiválasztani. Az ismertség ugyancsak nagyon fontos tényező; szükség van arra, hogy a hallgató jól ismerje az utánzott sze-

mélyt; elengedhetetlen, hogy beszédének akusztikai sajátosságai a hallgató „agyában” korábban már rögzítve legyenek. Mindennapi tapasztalat, hogy ahhoz, hogy valakit csupán a beszéde alapján felismerjünk, sokszor kell őt hallanunk, viszonylag gyakori egymásutánban.

Az elmondottak alapján feltételeztük azt, hogy az utánczó és az utánczott személy beszéde akusztikailag nagyon hasonló kell, hogy legyen. Ha ez a feltevésünk helyes, akkor meghatározhatók a beszédnek azok a komponensei, amelyek alapján a hallgató „becsapható”, vagyis az utánczás sikeres. Ez egyúttal azt a kérdést is felveti, hogy mi a viszony az utánczó saját beszédének és utánczós beszédének az akusztikai szerkezete között. A feltevések igazolására többféle kísérlet-sorozatot folytattunk le.

### **Anyag és módszer**

József Attila *Ars poetica* című versét választottuk hanganyagnak; amely egy jól ismert színész előadásában rendelkezésre állt, és amelyet egy ugyancsak ismert (színész) utánczó előadásában rögzítettünk. Az utánczó egyfelől az eredeti színészi előadást imitálta, a második felvételkor pedig saját személyében mondta el a verset. Ez a háromféle felvétel adta a kísérletek beszédanyagát. Az első vizsgálatsorozatban mindhárom anyagot akusztikailag elemeztük (az alaphangmagasság átlagát, alakulását, a beszédhangok időtartamát, frekvenciaszerkezetét, az egyes beszédszakaszok dallamstruktúráját, a hangsúlyozási megoldásokat, a beszédtempót), a CSL-4300-es digitális jelfeldolgozó különböző beállításaiiban. A vers 166 szóból áll; az elemzett beszédhangok száma 1200. A következő kísérletsorozatban pedig percepcióss teszteket végeztünk magyar anyanyelvű, az utánczott színészt ismerő személyekkel (ez utóbbiról a tesztelést követően kértünk információt).

A percepcióss teszteléshez a hanganyagot – laboratóriumi körülmények között – a lehallgattatási célnak megfelelően dolgoztuk fel és rögzítettük: a) változó hosszúságú soronként összekevertük az eredeti bemondást az utánczossal, majd az utánczó eredeti ejtésével is, b) az így létrehozott anyagot frekvenciaszűrőn engedték át (300 és 3000 Hz között) és c) különböző szótagszámú szavakat véletlenszerű sorrendben rögzítettünk először az eredeti és az utánczott, majd mind a három felvételből.



Az első kísérleti anyagot mintaadással és anélkül is lejátszottuk a résztvevőknek. A mintaadás azt jelentette, hogy az utánzandó színész ejtésében közöltünk két mondatot, kétszeri ismétléssel. A szűrt anyaghoz mintát nem adtunk. A szavak teszteléséhez a következőket választottuk ki: *költészet, csillaga, egymást, szegyet, mindenség, sziszegve, munkás, szertedüböörögni, paraszt*.

A kísérleti csoportoknak különböző feladatokat kellett megoldaniuk. Az a) esetben arról kellett – egyszeri lehallgatás alapján – dönteniük, hogy kit hallottak (minta nélkül). Majd azt kellett megítélniük, hogy az eredeti színészt vagy az utánzót hallották-e (egyik esetben sem közöltük a kísérleti személyekkel a beszélők nevét). Soronként kellett jelölniük az azonos beszélőket. A b) esetben is differenciálás volt a feladat. Ekkor a kísérleti személyek már tudták a színész és az utánzó nevét. A feladat megegyezett az előzőekben leírtakkal. A c) esetben ugyanazon szó két (eredeti és utánzott), majd három ejtéséhez (eredeti, utánzott és az utánzó saját ejtése) kellett a beszélő személyt hozzárendelni. A színész és az utánzó neve ismert volt, mintát azonban a résztvevők most sem hallottak. Az egyes csoportokat 15-15 fő alkotta (nők, férfiak vegyesen, az átlagéletkor 42 év volt). A teszteléseket egy hét különbséggel végeztük.

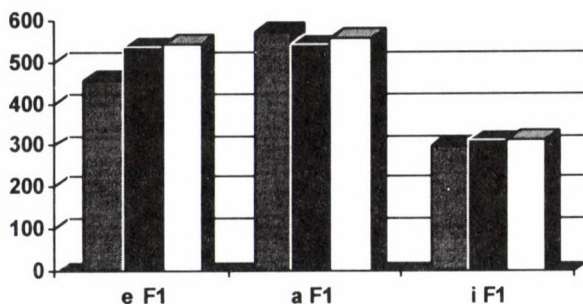
### Eredmények

Az első vizsgálatssorozat eredményeit az **akusztikai elemzések** adták. Összesítettük a beszédhangok formánsaira, az időtartamokra, az alaphangmagasság és az intenzitás változásaira kapott értékeket, majd – ahol szükséges volt – statisztikai elemzéseket is végeztünk. Az egyes beszédhangokra vonatkozóan az összes lehetséges előfordulást adatoltuk (az azonos kontextus követelményét ennek ellenére sem tűzhattük ki célul minden esetben).

A beszédhangok akusztikai szerkezetének elemzése sajátos különbségeket, illetőleg egyezéseket mutat, egyfelől a beszélőtől függően, másfelől aszerint, hogy az utánzó utánozni kíván-e avagy „saját maga” beszél. A magánhangzók jellegzetesebbek, a mássalhangzók artikulációjában jellegzetes eltérés alig volt található. A magánhangzók közül az [e, ɔ, i] hangok első formánsai a színész ejtésében egyértelműen más frekvenciasávban realizálódnak, mint az utánzó esetében. Kisebb

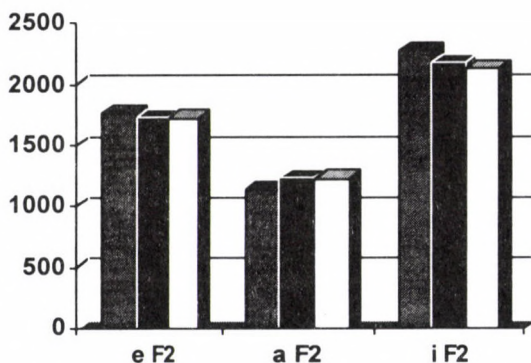


mértékű eltérések mutatkoznak a második formánsoknál is (1. és 2. ábra).



1. ábra

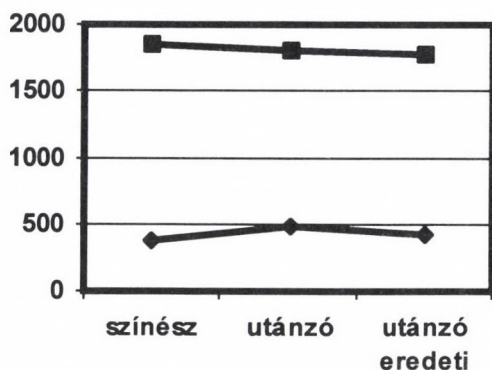
Az első formánsok alakulása a háromféle beszédanyagban (a szürke oszlop az [e], a fekete az [ɔ], a fehér az [i] átlagát szemlélteti)



2. ábra

Az első formánsok alakulása a háromféle beszédanyagban (a szürke oszlop az [e], a fekete az [ɔ], a fehér az [i] átlagát szemlélteti)

Az utánzó kétféle ejtésében alig van különbség a beszédhangok F1-ének és F2-ének átlagértéke között. Az [ø] magánhangzó különösen jellegzetesen alakul a háromféle ejtésben, ami azért sajátos, mivel akusztikailag ún. „semleges” magánhangzóként kevésbé feltételeztük, hogy jellegzetes különbség mutakozzon a formánsok átlagértékeiben. Az adatok kismértékű eltérést mutatnak az utánzó (utánzott és saját) anyagában, és relatíve nagyfokú különbséget regisztráltunk az utánzó és a színész eredeti ejtése között. A 3. ábra egy képre vetítve szemlélteti az első és a második formánsok átlagértékeinek alakulását.



3. ábra

Az [ø] magánhangzó első (◆—◆) és második formánsainak (■—■) átlaga a három beszédanyagban

A többi magánhangzó esetében nagy értékkülönbségeket a háromféle ejtésben nem találtunk; a tendencia azonban kivétel nélkül megerősítette a három kiemelt beszédhangra jellemző eredményeket. Az utánzó imitált és saját eredeti artikulációja hasonlóbb, mint amikor tudatosan igyekezett egy másik beszélőt utánozni. Lássunk néhány adatot! Az [a:] magánhangzó első formánsainak átlagértéke a színész, az utánzó

és az utánzó eredeti ejtésben: 662 Hz, 660 Hz és 675 Hz. A második formánsok átlagértéke ugyanezen sorrendben: 1320 Hz, 1354 Hz és 1393 Hz. Hasonló a helyzet például az [e:] magánhangzó esetében is. Az első formánsok átlagértékei (az előző sorrendben): 378 Hz, 395 Hz és 414 Hz, a második formánsoké pedig 2081 Hz, 2068 Hz és 2087 Hz.

Elemeztük a harmadik formánsokat valamennyi mérhető helyzetben, az összes magánhangzónál. A statisztikai elemzések szignifikáns eltérést a háromféle ejtés alapján nem mutattak, a legtöbb esetben a szórás meglehetősen nagy volt. Az átlagértékek a következők (az eltérések 20-40 Hz körüli értékek, amelyek az F3 esetén jelentéktelenek). Az F3 átlaga a színész ejtésében 2722,7 Hz, az utánzó ejtésében 2707,7 Hz, az utánzó saját ejtésében pedig 2683,2 Hz.

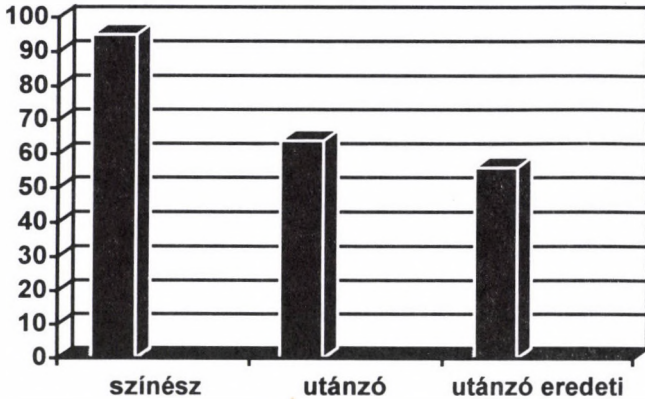
A színész jellegzetesen artikulálja az [e] magánhangzót, lényegesen zártabban, mint a köznyelvben szokásos hangzás. E sajátosan zártan ejtett [e] magánhangzók első és második formánsának értékeit összegeztük. Az F1 átlaga 456,2 Hz, a második formánsé 1760,3 Hz. A „zártág” akusztikailag – a köznyelvi ejtés frekvenciájához képest – az első formáns értékének csökkenésében jelentkezik. Az utánzó ezt a zárt ejtést igyekszik megvalósítani, az ennek a módosított artikulációnak megfelelő akusztikai paraméterek az alábbiak. Az első formánsok átlagértéke 532 Hz, a második formánsoké 1720 Hz. Ezek az adatok azt mutatják, hogy a zárt ejtést közelíti ugyan az utánzó, de nem sikerül minden esetben ugyanolyan mértékben realizálnia azt. Természetesen számos olyan [e] magánhangzót mértünk, amelynek az első két formánsa erősen megközelíti vagy eléri a színész ejtésének megfelelő formánsstruktúrát, de mégsem ezek tekinthetők gyakoriaknak.

Elemeztük a mássalhangzókat, amelyek artikulációjában mindössze egyetlen jellegzetes eltérést tapasztaltunk az ún. köznyelvi ejtémódtól. A színész hehezetesen artikulálja a zöngétlen zármássalhangzókat, különösen abszolút szó végén. Ez az akusztikai képen egyértelműen jelentkezik: a zárfelpattanást követően alacsony intenzitású zörejjelemek jelentkeznek csaknem a teljes spektrumban, amelyek időtartama jóval meghaladja a magyar ejtésben normatívnak tekinthető értékeket. Megnéztük, hogy ez az ejtési sajátság miként érvényesül az utánzó artikulációjában. Saját ejtésében egyáltalán nem jelentkezik,



vagyis a hehezetes zárhangokat nagyon tudatosan csak az imitált beszédben alkalmazza. Hangzásra, vagyis a hallgató észlelésében az eredeti és az utánzott mássalhangzók azonosnak tűnnek. Meglepetésre, az akusztikai szerkezetük jellegzetes eltérést mutat. Amíg a színész ejtésében a hehezetesség a fent leírt struktúrát mutatja, addig az utánzó a kívánt hangzást egy nagyon intenzív zárfelpattanással „helyettesíti”. Az artikulációs mozgás tehát teljesen különböző, a hangzásélmény azonban nagyon hasonló.

Elemeztük az egyes beszédhangok időtartamát a három szövegben. Az eredményeket a 4. ábra grafikonja szemlélteti.

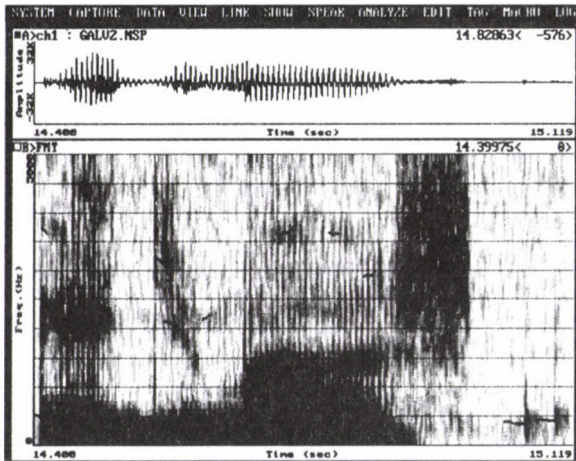


4. ábra

A beszédhangok átlagidőtartamának alakulása a három beszédanyagban

Az adatok szignifikáns eltérést mutattak ( $p < 0.01$  szinten) a színész és az utánzó ejtése között, amikor az utánzó a színész beszédprodukciónak imitálta. Ismételten nem találtunk ilyen mértékű különbséget az ugyanazon beszélő (azaz az utánzó) artikulációjakor.

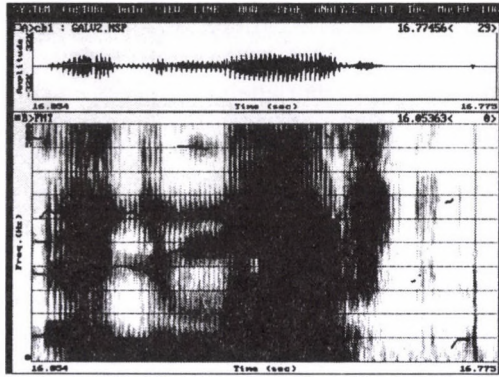
Az ugyanazon beszélő artikulációs sajátosságainak állandósága erősen relatív; nehéz ugyanis azt megmondani, hogy milyen mértékű különbségek jellemzőek még ugyanazon beszélőre és melyek már nem. Nincs rendszeres vizsgálati adatmennyiség arra nézve sem, hogy vajon mely tényezők hatnak jobban vagy kevésbé a beszéd artikulációjára, s mely akusztikai paraméterek értékeiben érhetőek e változások tetten. Vizsgálati anyagunkban – még a szándékolt eltérés, azaz az utánzás ellenére is – a szavak akusztikai szerkezete sokkal nagyobb hasonlóságot mutat ugyanazon beszélő ejtésében, mint az eredeti és az utánzott szavak esetében. A hangszíneképeken látható akusztikai szerkezet szemlélteti az eltéréseket és a hasonlóságokat (5., 6., 7. ábra).



5. ábra

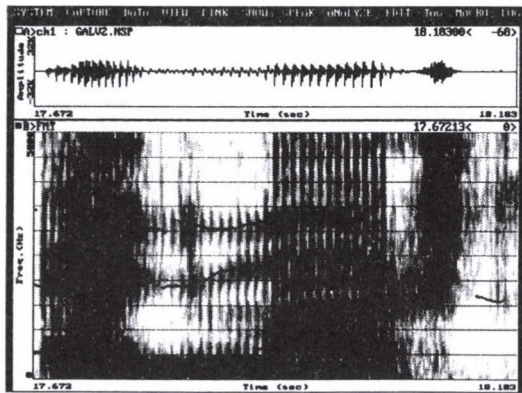
A színész ejtésében az *egymás* szó

Az akusztikai elemzések azt mutatják – példaként az *egymás* szót hoztuk az ábrákban –, hogy az utánzás akusztikai következményei csupán hasonlóak az eredeti ejtéshez.



6. ábra

Az utánzó ejtésében az *egymás* szó (utánzasként)



7. ábra

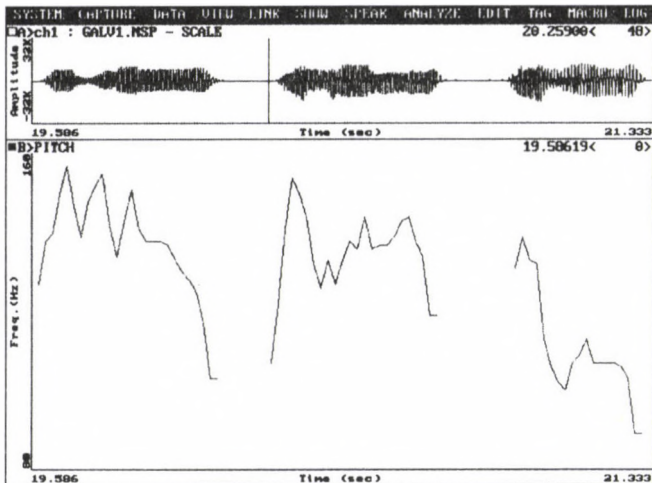
Az *egymás* szó ejtése alapján készült spektrogram az utánzó saját ejtésében

Összegzésként kimondható, hogy a beszéd akusztikai szerkezetének vizsgálata semelyik területen sem mutatott ki olyan jellegű egye-



zéseket, amelyek a sikeres utánzást, illetőleg a hallgató benyomását kétséget kizáróan igazolták volna.

Vizsgáltuk az alaphangmagasság határértékeit, valamint a beszéddallamok alakulását. A színész alaphangmagasságának határértékei 107-155 Hz, az utánzóé pedig 103-140 Hz, vagyis az utánzó alaphangja kissé mélyebb. A színész utánzásakor az utánzó az alaphangmagasságát jellegzetesen megemeli, tehát közelít az utánzott beszélő értékeihez. Az utánzott beszéd F0-határértékei 125-150 Hz szórnak. Az utánzó igyekszik továbbá az eredeti beszélő jellegzetes dallamvonulatait reprodukálni. A 8. ábra mindháromra hoz példát (ugyanazon mondat esetében).



8. ábra

A színész (az eredeti), az utánzott és az utánzó saját ejtése alapján készült dallamgörbék (balról jobbra)

Elemeztük az intenzitásváltozásokat. A színészre jellemző, hogy nagy különbségekkel realizálja az egyes beszédszakaszokat. Az utánzó az imitálás során ezt a hangerőstruktúrát igyekszik megvalósí-

tani. Helyenként túlzottan is él az intenzitás adta lehetőségekkel, nagyobb különbségeket hoz létre, mint amilyenek az eredeti beszélőnél tapasztalhatók.

A **percepciók eredmények** magyarázattal szolgálnak arra, hogy miért állítottuk az akusztikai elemzések után, hogy nem kaptunk egyértelmű választ a hangzásbeli hasonlóságra. Látszólagos ellentmondást tapasztaltunk ugyanis az objektív adatok és a szubjektív benyomás között. A beszédutánzás sikere ugyanis több tényező, illetőleg azok együttjárásának függvénye.

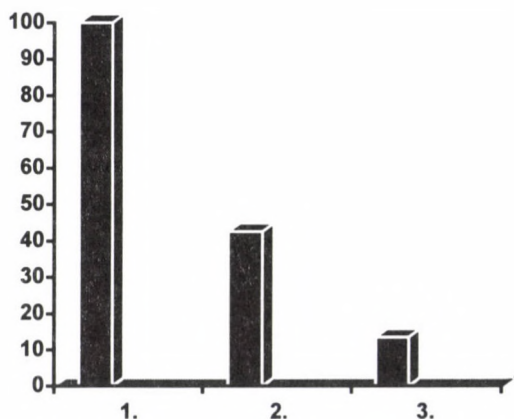
(i) Tökéletesen sikeres akkor, ha nincs azonos időben „versengő” beszédminta, vagyis ha az eredeti beszéd nem hangzik el az utánzást közvetlenül megelőzően. Ha a kísérleti személyek csupán az utánzót hallották és azonosítaniuk kellett a beszélő személyt, akkor az eredeti színészt 98,8%-ban ismerték fel. Az utánzó tehát képes volt a hallgatók „megtévesztésére”, ha aktuális összehasonlítás nem történhetett.

(ii) Ha csak differenciálni kellett az eredeti és a „másolat” között, vagyis rendelkezésre állt valamiféle „élő” minta is, akkor a hallgatók teljesítménye szignifikánsan javult, azaz az utánzás sikere szignifikánsan gyengült.

(iii) Az utánzás sikere nem volt független a hallott szöveg hosszától sem. Az eredeti beszélő biztonságos felismerése annál pontosabb, minél több beszédrészlet áll a hallgató rendelkezésére ahhoz, hogy az agyban tárolt „adatokat” az adott hangzásélménnyel összevesse. Minél több, illetőleg minél hosszabb a meghallgatott beszédrészlet, annál egyszerűbb a döntés. Kísérletünkben a szavak álltak szembe a mintegy mondatnyi hosszúságú beszéddel. Amikor a beszélőket a hallgatóknak szavak alapján kellett megkülönböztetniük, 43,4%-os téves eredményt kaptunk. A hibaarány 12%-ra csökkent, ha rövid szakaszok differenciálása volt a feladat.

Különösen nehéznek bizonyult a feladat, ha szavak alapján kellett a kísérleti személyeknek a három ejtést elkülöníteniük (a színészt, az utánzót és az utánzó eredeti szavait). Ebben a kísérletben közöltük ugyan a hallgatókkal, hogy ugyanazokat a szavakat fogják hallani egymást követően, de különböző beszélők kiejtésében. A szó elhangzása után azonnal dönteniük kellett a beszélő személyről. Aktiválniuk kellett az agyukban a két beszélőhöz kapcsolódó neurális spektrogramo-

kat, az aktuális felidézést segítő minta nélkül. Az eredmények határozott romlást mutatnak az előző kísérleti eredményekhez képest. A szavak azonosítása mindössze 43,4%-ban sikeres (ez majdnem véletlenszerű találatnak tekinthető érték). A résztvevők nemcsak a színészt nem tudják pontosan azonosítani, de nemegyszer az utánzó saját hangján ejtett szót sem ismerik fel helyesen. A két-két hallott szó közötti differenciálás lényegesen könnyebb feladat volt, mint a három szó közötti (9. ábra).



9. ábra

A téves azonosítások aránya a minta jelenléte és a tesztanyag nagyságától függően (1.= azonosítás eredeti minta nélkül, 2.= szavak azonosítása mintával és 3.= szöveg azonosítása mintával)

Azok a kísérleti személyek, akik nem emlékeztek nagyon jól az eredeti színészre vagy ritkábban látták, hallották, érthetően gyengébb teljesítményt nyújtottak (ők főleg a fiatalabb korosztályokból kerültek ki). Az idősebbek teljesítménye lényegesen jobb volt, a színész be-



szédéről tárolt „engram” sokkal pontosabb képet rögzített az összevetésekhez.

A szűrt anyaggal végzett kísérlet eredményei csaknem azonosak voltak az eredetiekével. Az egyedüli különbség abban jelentkezett, hogy a színészt kevésbé ismerők jobban „becsaphatók” voltak, azaz az utánzás náluk sikeresebben működött.

### **Következtetések**

Az e tanulmányban leírt kísérletsorozat egy széleskörű kutatás része, amelyben az egyénre jellemző fonetikai sajátosságokat kívánjuk meghatározni, leírni. A jelen kísérlet eredményei módszertani szempontból is jelentősek, hiszen egy egészen más aspektusból közelítik az egyéni hangszínezet megismerhetőségét. Az eredmények azt mutatják, hogy a beszéd utánosztható és a hallgató „átejthető”, vagyis elhitethető vele, hogy az utánzott személyt hallja. Adataink azonban egyértelműen rávilágítottak arra is, hogy az utánzás sikere korlátozott, több tényezőtől függ és rendkívül változékony.

Az utánzó a saját artikulációs mozgásait igyekszik hasonlatossá tenni az utánzott személy ejtémódjához. Nem az adott artikulációs gesztusoknak van döntő jelentősége, hanem az akusztikus következménynek, még pontosabban annak a perцепciós élménynek, amit az a hallgatóban kivált. Már Kempelen Farkas is leírta, hogy a fülünk „becsapható” (1871), hiszen beszélőgépeinek létrehozásakor ezt a tényt tudatosan fel is használta.

Az utánzó tehát egyfelől közelíti saját ejtését az utánzott személyéhez, másfelől pedig mintegy felnagyítja, illetőleg még jellegzetesebbé teszi az utánzottra jellemző ejtési sajátosságokat. E kettő kombinációjával éri el a kívánt hatást: nem ő, hanem valaki más, az a bizonyos személy beszél. Hasonlatos mindez a karikatúrához. A jó karikatúra hasonlít ugyan az eredeti személyhez, annak egyértelmű azonosítását azonban döntően az eredeti személy jellegzetességeinek (nemritkán eltúlzott) vizuális hangsúlyozásával éri el.

### Irodalom

Coulthard, M.: Forensic discourse analysis. In: *Advances in Spoken Discourse Analysis*. Routledge. Ed.: Coulthard, M. London 1992, 242-258.

Gósy Mária–Nikléczy Péter: A beszélő felismerése a beszéde alapján: elméleti háttér és módszertani megközelítések. In: *Beszéd kutatás '99*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1999, 1-20.

Kempelen, W. von: *Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine*. Wien 1871.

Schlichting, F.–Sullivan, K.P.H.: Can voice imitation be detected in voice line-ups in a language unknown by the listeners? *Phonum* 6. 1998, 105-118.

A kutatás a T 025965 sz. OTKA-munkálat keretében folyt.

# **BESZÉDADATBÁZISOK KÉSZÍTÉSE GÉPI BESZÉDELŐÁLLÍTÁSHOZ**

**Olaszy Gábor**  
**MTA Nyelvtudományi Intézete**

## **Bevezetés**

A beszédatadbázis viszonylag új fogalom, a digitális beszédfeldolgozás fejlődése hívta életre. Jelen munkában a beszédatadbázis szóval olyan adatbázist jelölünk, amelyből a gépi beszédelőállítás során felépítjük a kimondandó beszédjelet. Tehát ez a beszédatadbázis tartalmazza a szintetizálendő beszédjel létrehozásához szükséges hullámforma jelrészleteket. Az adatbázis – a szintézis és az alkalmazás fajtájától függően – lehet nagyon egyszerű is, és lehet nagyon bonyolult. Az adatbázisban elhelyezett jelrészletek minőségétől alapvetően függ az előállított beszéd minősége. Jelen tanulmányban összefoglaljuk, hogy milyen fonetikai szempontokat célszerű figyelembe venni a beszédatadbázisok tervezésénél, készítésénél. A megállapításainkat csak emberi beszéd alapú, hullámforma elemekre tesszük, nem foglalkozunk a formáns alapú beszédatadbázisok problémakörével.

## **1. A beszédatadbázisok fajtái**

Alapvetően háromféle adatbázissal foglalkozunk, az első, amelyik egyedi beszédelemeket tartalmaz, a második az úgynevezett diádus felépítésű elemeket tartalmazó, amelyben két hang kapcsolódása jelent egy-egy elemet, a harmadik pedig a kevert szerkezetű adatbázis, amelyben diádus, triádus (három hangelem), és esetlegesen a triádnál is hosszabb elemek találhatók. A beszédatadbázis elkészítése minden esetben adott (előre megtervezett) szöveg felolvasásával és hangrögzítéssel történik. Az így kapott beszédjelből készítik el az adatbázis elemeket.



**Egyedi beszédelemeket** olyan adatbázisoknál alkalmaznak, amelyek ún. kötött szótáras rendszereket szolgáltatnak ki. Kötött szótáras rendszernek nevezzük azokat a rendszereket, amelyekben a szintetizálendő beszédüzeneteket előre meghatározzák és azokat olyan elemekből állítják össze, amelyek viszonylag hosszabb beszédszakaszokból állnak (szó, mondatrész, mondat). Ilyen rendszerek működnek ma sok telefonos szolgáltatásban (ébresztő szolgálat, tudakozó, számváltozás bemondó, hangos számla szolgáltatás, hangposta, MÁV pályaudvari utastájékoztató, telebank szolgáltatások, hangos hirdetések stb.). Az elemek száma az ilyen adatbázisokban az alkalmazástól függően lehet nagyon kevés vagy akár pár ezer is. Az egyedi beszédelemek megtervezése és elkészítése kis elemszám esetén egyszerű, azonban az elemszám növekedésével egyre bonyolultabbá válik. Az 50 elemnél többet tartalmazó kötött szótáras rendszerek (például: hangposta) beszédelemeinek jó minőségű elkészítése már igen gondos tervezést, szakszerű hangfelvételt kíván.

**Előnyök:** az ilyen adatbázisok előnyös tulajdonsága, hogy az üzenetek nagyon jó minőségben szólalhatnak meg, kis elemszám esetén az elkészítésük nem kíván fonetikai szakértelmet.

**Hátrányok:** nagyobb elemszám esetén fonetikai, beszédakusztikai szakismeretek szükségesek a tervezéshez. Ellenkező esetben az adatbázis elemeiből elkészített beszédüzenet kifejezetten rossz hangminőségű is lehet.

**A kizárólag diádos elemeket** (két félhangot) tartalmazó adatbázisok általában a kötetlenül beszélő, ún. szövegfelolvasó rendszerekben használatosak. Ezekre az adatbázisokra az jellemző, hogy a bennük eltárolt beszédjel elemeket úgy alakítják ki, hogy a beszédhangokból két hang kapcsolódását tárolják el, ezt is úgy, hogy az elemhatárokat a két hang közepén jelölik ki. Az ilyen elemek tehát két félhangot tartalmaznak. A *báb* hangsort tehát a *#b*, *bá*, *áb*, *b#* elemekből lehet felépíteni (a *#* jel a csendes szakaszt jelöli). Az elemek száma az ilyen adatbázisokban általában ezres nagyságrendű. Ez a szám a nyelvtől, azon belül pedig a szintézishez meghatározott beszédhangok számától függ. A szintézishez meghatározandó beszédhangok száma nem egy-

értelmű, a tervező döntésének a függvénye. Vegyük például a magyar beszédet. A magánhangzók számát meghatározhatjuk 9-nek, de lehet 14 is, ha ezekhez hozzávesszük a rövid magánhangzók hosszú párjait is. Ugyanis nem feltétlenül szükséges, hogy a hosszú magánhangzókat külön hangként kezeljük, azok előállíthatók a rövid párjuk felhasználásával nyújtással is, csak kissé szokatlanabb hangminőséggel. Ez azonban a megértést nem befolyásolja. A magánhangzók listáját még tovább bővíthetjük azok nazalizált változataival, amelyek bizonyos hangkapcsolatokban (*ank, enk, onk, ing, eng, áng* stb.) törvényszerűen kialakulnak. A mássalhangzók tekintetében is többnyire egyértelmű, hogy mely hangoknak kell szerepelni a beszédhangok listájában, de lehetnek olyanok is, amelyeket például csak egy igen igényes, kifogástalan beszédminőséget megcélzó rendszerben kell felvenni a listára. Ilyen például az [n] speciális változata, amely az [ŋg, ŋk] hangkapcsolatok hatására jön létre (például az *ing, engem, fánk, fiunk, tönk, ankét* stb. szavakban), de ilyenek a [dz, dʒ] hangok is avagy a [j] hang zöngétlen változata (például *kapj, lépj, lopj*).

**Előnyök:** a diados adatbázissal gondos tervezés és sok akusztikai csiszolás (az elemek egymással való kapcsolódási pontjainak precíz akusztikai illesztése) esetén alapvetően jó minőségű beszédet lehet előállítani. Ezt alátámasztja az is, hogy a világon általánosan használt ez a technológia.

**Hátrányok:** az adatbázis elkészítése igen munkaigényes. A hangfelvétel szövegének megtervezése és magának a hangfelvételnek az elkészítése gondos előkészítést, jól felkészült bemondót igényel. A digitális feldolgozáshoz bonyolult fejlesztői és vizuális megjelenítő rendszer szükséges, ami kereskedelmi forgalomban nem kapható (Németh et al. 1997). Akusztikai szempontból hátrány, hogy az adatbázis szerkezetéből adódóan a magánhangzók közepén vágási pontok vannak. Ez egyrészt azt eredményezi, hogy a formánsmenetek a magánhangzóban a vágási pontnál megtörhetnek, tehát a spektrális folyamatosság nincs biztosítva, másrészt a magánhangzóban megvalósított alaphérfrekvencia a két félmagánhangzóban lényegesen (5-15 Hz-cel) is eltérhet. Mindkét tényező enyhe torzítást eredményez a



hangban, ami a beszéd tiszta hangzását rontja. E hátrányok enyhítésére speciális simító algoritmusokat kell kidolgozni. Mindezek ellenére a diádós adatbázissal működő beszédelőállító rendszerek hangja tartalmaz torzításokat, ami azt eredményezi, hogy a hang nem teljesen tiszta, hanem alatta sustorgásra emlékeztető mellékörejek is vannak.

**A kevert felépítésű beszédatadbázisok** diádós, triádós és esetlegesen a triádnál is hosszabb elemeket tartalmaznak. Ezeket az adatbázisokat a kifejezetten jó minőségű, szöveg-beszéd átalakítókban alkalmazzák (Hirokawa et al. 1992). Az eldöntendő kérdés a tervezésénél az, hogy mely hangcsoportokat tároljanak diádós és melyeket triádós, illetve annál hosszabb formában. Az általunk kidolgozott magyar beszédatadbázisban a CC, VV, CV és VC elemek diádós, a CVC, elemek triádós, formában vannak tárolva. Az ilyen adatbázisoknál az elemszám ugrásszerűen megnövekszik, hiszen a diádós elemeken felül még el kell készíteni a triádós felépítésű elemeket is, vagyis a CVC struktúrát minden hangkapcsolatra vonatkoztatni kell. Például a C\*-V-C kapcsolatokat nézve (ahol a C\* például a [b] hang, C pedig a [p, d, t, l] stb. hang) egyértelmű, hogy a C\*-gal minden V és minden C kapcsolódást létre kell hozni az adatbázisban.

Példa a C\*VC elemekre, ahol C először a [p], majd a [d] mássalhangzó:

*báp, bap, bop, bóp, bup, búp, büp, búp, bip, bíp, bép, böp, böp, bep, bád, bad, bod, bód, bud, búd, бүd, būd, bid, bíd, béd, böd, böd, bed,*

Ez azt jelenti, hogy az ezres nagyságrendű diádelemeken felül még több ezer triád elemet is el kell készíteni.

**Előnyök:** a kevert szerkezetű adatbázis alkalmazásával igen jó minőségű, professzionális beszédelőállítót lehet készíteni. A magánhangzók formánsmenetei nem törnek meg, a folyamatos spektrális változás biztosított a magánhangzóban, és az alaphangban sincs törés, így a diádós adatstruktúrából adódó torzítások jelentősen csökkenthetők, ami – összehasonlítva a diádós adatbázisból készített beszéddel – határozottan tisztább hangot eredményez.



**Hátrányok:** a kevert adatbázis elkészítése sokkal munkaigényesebb, mint a diádosé, speciális fejlesztői és vizuális megjelenítő rendszert igényel. Memóriaigénye igen nagy (10-30 Mbyte).

## **2. A beszédatadbázisok akusztikai tervezése**

Általában elmondható, hogy a beszédatadbázis elemeinek akusztikai tervezése kihat az egész működő rendszer szolgáltatásának minőségére. Ezért fontos, hogy a tervezés fázisában már globálisan átlássuk az egész majdani beszélő rendszer működését, a vele szemben támasztott követelményeket. Külön tárgyaljuk az egyedi beszédelemeket tartalmazó adatbázisok (kötött szótárak) és külön a felolvasó rendszerekhez készített építőkockaszerű, hangelemeket tartalmazó adatbázisok tervezését.

### **2.1 Kötött szótárak beszédelemeinek tervezése**

Az egyedi, előre megtervezett üzeneteket megszólaltató kötött szótáras rendszerek beszédatadbázisainak akusztikai tervezése már az üzenetek szövegének meghatározásánál kezdődik. Általában kétfajta szöveges listát kell elkészíteni. Az egyik az a szöveg (szó, mondatrész, mondat), amely a majdan elhangzó „üzenet” elemeit tartalmazza. Ezeket a szövegelemeket a rendszer majd beszédhullám formájában fogja tárolni, és ezekből kell összeállítani a bemondandó teljes üzenetet. A másik írott anyag tartalmazza azt a „szöveg” korpuszt, amelyet a hangfelvétel elkészítésénél a bemondóval felolvastatunk. Célszerű, hogy a „szöveg” anyaga bővebb legyen, mint az „üzenet”-é. Sok mai beszéddel válaszoló rendszer hangminősége azért kifogásolható, mert a bemondáskor az „üzenetek” szövegét olvastatták fel a bemondóval. A megszólaló „üzenet” kétféle lehet egy kötött szótáras rendszerben: teljes mondat (M), és a mondatnál rövidebb beszédrészlet (BR). A teljes mondatot tartalmazó üzenetek (például: *A telefonszám megváltozott, kérjük hívja a tudakozót.*) mindig jó minőségben szólalnak meg, hiszen azok akusztikai szerkezetét a bemondó hangja, előadásmódja határozza meg. Ha a rendszer csak ilyen fix üzeneteket tartalmaz, akkor az „üzenet” szöveglista és a felolvassandó „szöveg” megegyezik, akusztikai tervezés nem szükséges. Ha

az üzenetek felépítéséhez nemcsak teljes mondatok, hanem beszéd-részetek felvétele is szükséges, akkor már szükség van akusztikai tervezésre. Ha az információ változik az üzenetben (például: *Az Ön postafiókjába ..3..... üzenet érkezett.*), akkor több BR elem összekapcsolásával hozzuk létre a bemondandó üzenetet. Az előbbi példamon- datban az üzenet első része (BR1) és vége (BR2) közé kell bevágni az aktuális számelemet mint változó információt. Az ilyen típusú mon- datok hangzásában – ha nincs akusztikai tervezés – általában meg lehet hallani, hogy az üzenetet több részből vágták össze. Három hibakategória határozható itt meg: a szaggatottság, továbbá a hangerő és a intonáció nem megfelelő volta. Mindhárom hiba mértéke csök- kenthető, ha a felolvasandó szöveg tervezésénél gondosan járunk el és figyelembe vesszük a beszédképzés fonetikai szabályait is. Ez azt je- lenti, hogy a felolvasandó „szöveg”-et úgy kell megtervezni, hogy a BR elemek a kiejtésüknek megfelelő szöveggörnyezetbe legyenek be- ágyazva és így kell a lista elemeit felolvasatni. Az 1. táblázatban né- hány példát adunk az „üzenet” és a „szöveg” kapcsolatára.

1. táblázat: Példák a szövegtervezésre

A rendszer „üzenet”- elemei	Tí- pus	A bemondandó „szöveg”, amely tartalmazza az üzenetelemet
1. Önnek ...	BR	1. <b>Önnek</b> hét <b>üzenete van.</b>
2. .... üzenete van.	BR	
3. Kérjük, nyomja meg a .	BR	2. <b>Kérjük, nyomja meg a csillag gombot</b>
4. ....gombot.	BR	
5. ....egy...	BR	3. Önhöz <b>egy</b> üzenet érkezett
6. ... kettő ...	BR	4. Önhöz <b>kettő</b> üzenet érkezett
7. ....csillag....	BR	5. Nyomja meg a <b>csillag</b> gombot
8. Önnek nincs több üzenete	M	6. <b>Önnek nincs több üzenete</b>
9. Az Ön üzenete elmen- tésre került.	M	7. <b>Az Ön üzenete elmentésre került</b>

Az „üzenet” elemeket (a kiemelt részek) a bemondott „szöveg” ele- mekből kell kivágni. Ezzel, mivel természetes környezetéből vágtuk ki az elemeket, biztosíthatjuk azt, hogy az összefűzésnél nem alakul-



nak ki olyan hibák az akusztikai szerkezetet illetően, amelyeket korábban említettünk. Fontos szempont a „szöveg” megtervezésénél, hogy legyünk tekintettel a kivágás megvalósíthatóságára is. Ez annyit jelent, hogy a szöveget úgy kell megtervezni, hogy a kivágandó elem határain lévő hangok a legkevésbé torzuljanak az előttük lévő, illetve az őket követő hang hatására. Ezért választottuk például az 1. szöveg két kivágandó eleme közé a *hét* szót és nem például a *kettő* szót. A *hét* szó utolsó hangja és az *üzenet* szó első hangjának kezdete is jól meghatározható. Ha a *hét* helyett a *kettő* szót tettük volna bele ebbe a mondatba, akkor a *kettő* *üzenet* hullámformájából az *őü* hangkapcsolatból sokkal nehezebben lehetne meghatározni az *üzenet* szó kezdetét, mivel magánhangzó találkozásban a hullámforma folyamatosan alakul át az egyik hangból a másikba. Más finomság is figyelembe vehető némi akusztikai tervezéssel. Például a *csillag gombot* üzenetrészt két elemből (a *csillag* és a *gombot*) kell összeállítani. Mivel a két elem határán ugyanaz a hang van, a zárfelpattanást nem kell megvalósítani mindkettőben. Ez a természetes artikuláció következménye. Ezért célszerű a *csillag* elem kivágásánál a zárfelpattanást elhagyni a [g] hangból, még akkor is, ha a bemondó ejtette. Az így elkészített elemmel való összekapcsolás után a *csillag gombot* beszédrészlet természetesen és folyamatosan fog hangzani, nem pedig mesterkélten, külön ejtve a két [g] hangot, hogy *csillag gombot*. Az ilyen fonetikai tervezés sokat javít az elemekből összefűzött beszéd minőségén. Természetesen a fonetikai tervezéskor alkalmazott szabályok a szövegtől függnek, tehát minden rendszerhez más és más szabályokat kell alkalmazni.

Az akusztikai tervezéshez hozzátartozik az is, hogy a megszólaló rendszer beszédüzeneteinek kialakításánál a lehető legjobban biztosítsuk a természetes ejtésre jellemző folyamatosan változó (törések, ugrások nélküli) spektrumképet. Ugyanez vonatkozik a prozódiai szerkezetekre is (hangintenzitás, dallam, időszerkezet). Azoknál a rendszereknél, ahol egy-egy üzenetet sok (4-10) részegység összekapcsolásával kell előállítani (például: számfelolvasás) különösen fontos a fentiek figyelembe vétele. Ezen követelmények teljesítéséhez ismerni kell



a csatlakozó elemek csatlakozó hangjainak (adott elem utolsó és a hozzá csatlakozó elem első hangja) spektrális szerkezetét és összekapcsolódásuk spektrumképét. (Olaszy 1989). A 2. és 3. táblázatban megadjuk azokat az általános fonetikai szabályokat, amelyeket a korrekt hangcsatlakozás érdekében érdemes figyelembe venni a szótárelemek tervezésénél. A táblázatokban megadott szabályok általános érvényűek (a példákban egy számfelolvasó elemeire való alkalmazást szerepeltettük). A táblázatok szabályainak figyelembe vétele azzal jár, hogy egy kötött szótáras rendszerben egyazon elemből többféle variánst kell eltárolnunk és ezekből kell az összekapcsolásnál az erre kialakított speciális algoritmussal kiválasztani az egymáshoz fűzendő elemeket.

2. táblázat: A hangkapcsolatok egymásra hatása (visszafelé)

Az előző elem eredeti utolsó hangja (betűképpel)	Az előző elem megváltozott utolsó hangja, ha	a hozzá csatlakozó elem első hangja	Példa
<i>b, d, g, v, z, sz</i>	<i>p, t, k, f, sz, s</i>	<i>p,t,k,c,cs,sz,s,f,h</i>	
<i>c</i>	<i>c</i> zárfelpattanás nélkül	<i>sz</i>	<i>Kilencszáz</i>
<i>t</i>	<i>t</i> zárfelpattanás nélkül	<i>n</i>	
<i>n</i>	<i>n(k)</i>	<i>k</i>	<i>Tizenkettő</i>
<i>n</i>	<i>n(g)</i>	<i>g</i>	
<i>n</i>	<i>n(h)</i>	<i>h</i>	<i>Huszonhárom</i>
<i>n</i>	<i>nn</i>	<i>n</i>	<i>Ötvennégy</i>
<i>n</i>	<i>ny</i>	<i>ny</i>	<i>Hatvannyolc</i>
<i>n</i>	<i>m</i>	<i>m, b, p</i>	<i>Ötvenmillió</i>
Magánhangzó	magánh.+átmenet	magánhangzó	
Magánhangzó	magánh.+átmenet	<i>gy,ty,ny</i>	

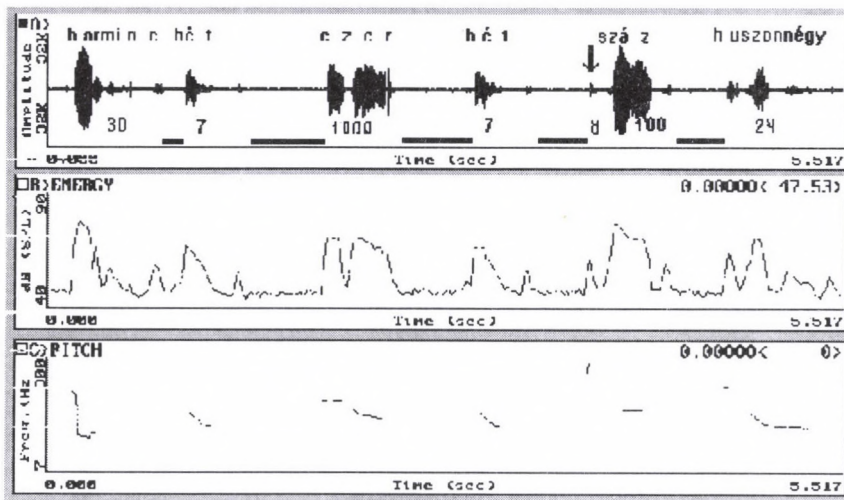
Az oszcillogramból (A) látható, hogy a beszédelemeken kívül zajok, pattogások is vannak a hangban. Sőt, a „száz” elem kivágásánál az elem előtt egy „sva”-szerű rövid hangrészt is benne hagyott (a nyíl mutatja) a rendszert készítő. Ebből látszik, hogy a banki kötött szó-

táras rendszer elkészítése során felületesen dolgozták ki a számelemek hangfájljait.

### 3. táblázat: A hangkapcsolatok egymásra hatása (előre)

Az előző elem utolsó hangja (betűképpel), ami előre hat a következő hangra	A csatlakozó elem eredeti első Hangja	A csatlakozó elem megváltozott első hangja	Példa
N	Magánhangzó	nazalizált magánh.	ötvenezer
ny,gy,ty	Magánhangzó	átmenet + magánh.	négyezer
Magánhangzó	Magánhangzó	átmenet + magánh.	kettőezer

A rossz, gondatlan tervezés példaként bemutatjuk, hogy egy mai banki telefonos számlafelolvasó szolgáltatásban a gép hogyan olvassa fel a 37 724 számot (1. ábra).



1. ábra

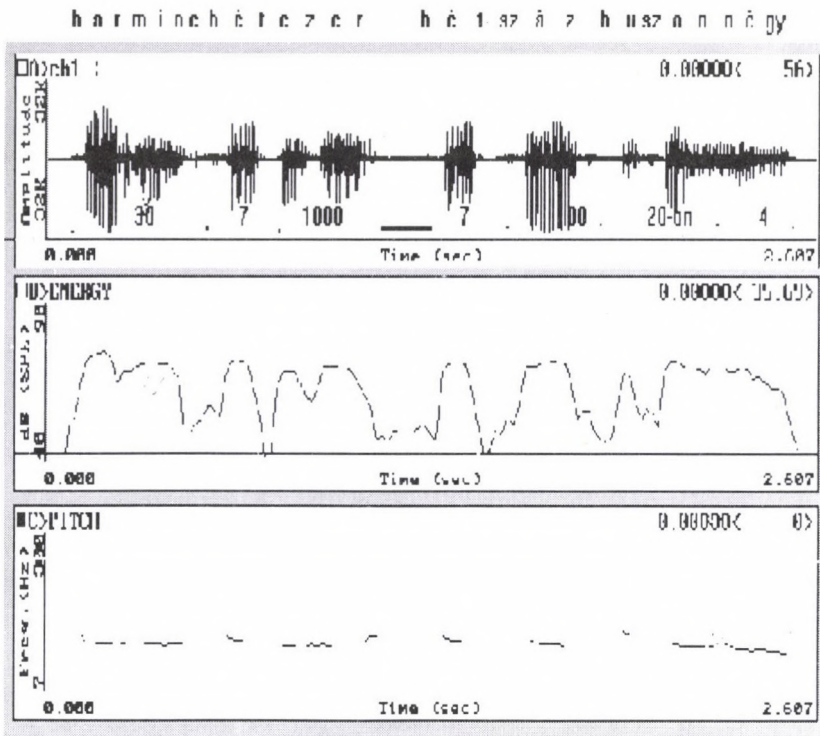
A 37 724 szám kimondása egy banki rendszerben

Az oszcillogramon továbbá látható, hogy nagyon szaggatottan hangzik fel a szám. Az elemek közé 150-500 ms-os szüneteket iktattak a fejlesztők (ezzel próbálták azt kompenzálni, hogy az adott szám-elemek egymáshoz való illesztése nincs megvalósítva). A szüneteket vízszintes vonallal jelöltük az ábrán. Ez a természetellenes megoldás döcögössé, érthetetlenné teszi a számot. A B ablakban a kimondott szám hangintenzitás-görbéjét ábrázoltuk. Látható, hogy az „ezer” és a „száz” elemek mintegy 6-10 dB-lel intenzívebben szólnak, mint például a „hét”. Ez szintén természetellenes, mivel a számok kimondásakor a helyiértéket megtestesítő részek mindig hangsúlytalanok, tehát intenzitásuk semmiképpen nem lehet nagyobb, mint a hozzájuk kapcsolódó elemé. A C ablakban a szám dallamgörbéjét ábrázoltuk. Látható, hogy a „harminc” elem hanglejtése ereszkedő, mintha befejező elem lenne, noha ez az első eleme a számnak. Az „ezer” és a „száz” elemek hangmagassága magasabb, mint a „hét” elemé, ami a természetes ejtésben sohasem fordul elő. Mindezek a hibák összeadódnak a kiejtés során és ennek eredménye egy igen furcsa számkimondás.

A 2. ábrán bemutatjuk, hogy ugyanezt a számot hogyan olvassa fel egy gondos akusztikai tervezéssel készített rendszer. Az ablakok itt is ugyanazokat az adatokat tartalmazzák, mint az 1. ábrán. Az oszcillogramból látható, hogy csak ott van szünet, ahol a szám kiejtési logikája megkívánja (az „ezer” után), a többi rész folyamatosan hangzik el.

Ez látszik az összidőtartamon is, itt a szám kimondása mindössze 2,6 másodperc, míg az 1. ábrán látható kimondásban ugyanezen szám kimondásához 5,5 mp-et használt fel a rendszer a beiktatott szünetek miatt. A B ablakban látható, hogy az intenzitásvonal kiegyenlített és enyhe csökkenést mutat, ahogy az a normál beszédben is megszokott. A C ablakban pedig a dallamgörbét láthatjuk, amely szintén kiegyenlített vonulatot mutat, tehát megfelel a normál beszéd dallamának.





2. ábra

A 37 724 szám kimondása a gondos akusztikai tervezéssel készített számfelolvasóval

## 2.2 Adatbázis-elemek tervezése automatikus felolvasáshoz

Az automatikus szövegfelolvasáshoz készített diádos beszédat-bázis tervezése lényegesen bonyolultabb megoldásokat kíván, mint amilyeneket kötött szótáras rendszereknél alkalmaznak. Ezt mutatja az is, hogy mind a mai napig egyetlen olyan cég sem kínál ilyen szövegfelolvasót, amelyik egyébként kötött szótáras rendszereket készít. Ennek az az oka, hogy míg a kötött szótáras rendszerek tervezéséhez

nélkülözni lehet a nyelvészeti-fonetikai szakértelmet (bár ez, mint láttuk az 1. ábrán, a minőség rovására mehet), addig az automatikus szövegfelolvasó rendszerek tervezéséhez erre mindenképpen szükség van. Nem véletlen tehát, hogy az első szöveg-beszéd átalakítót Magyarországon (HUNGAROVox) az MTA Nyelvtudományi Intézetének Fonetikai Laboratóriumában fejlesztették ki a 80-as évek elején.

### 2.2.1 Diádos adatbázis-elemek tervezése

A diádos rendszerű adatbázis tervezésének első lépése a beszédhangok állományának meghatározása, vagyis annak eldöntése, hogy mely beszédhangok fognak megszólalni a rendszerben. A beszédhangok többsége az adott nyelv ismeretében általánosan meghatározható, azonban lehetnek olyan hangok is, amelyekről külön dönteni kell. Ilyen hangok a magyarban például a nazalizált magánhangzók (pl. *inga, engem, hangos, bankár*), a [j] hang zöngétlen változata (*kapj, lépj, lopj*), a [h] hang zöngés változata (*nahát*). Ide tartoznak a ritkán előforduló beszédhangok is, mint például a [dz, dʒ] és ezek hosszú változatai. De ebbe a kategóriába kell sorolni a hosszú magánhangzókat is, amelyeknél azt kell eldönteni, hogy megelégszünk-e azzal, hogy ezeket a hangokat a rövid párjuk megnyújtásával fogjuk majd előállítani, vagy pedig mint önálló hosszú hangként fogjuk őket kezelni. Ez a döntés több tényezőtől függhet. Egyrészt látni kell, hogy ha növeljük a kezelendő beszédhangok számát, akkor növekszik a diádos adatbázis mérete, bonyolódik az adatbázist felhasználó keretrendszer szerkezete, másrészt azt is tudni kell, hogy minél precízebben határozzuk meg a hangkészletet, annál pontosabban tudjuk a későbbi beszédjelet előállítani, ami a jó minőséget garantálja. A magyar beszéd hangkészletét beszédatadbázis létrehozása céljából célszerű 14 magánhangzóból és 24 mássalhangzóból meghatározni. Ehhez még hozzá kell venni a szünet elemet is, amit a hangkezdő és hangvégződési helyzetű elemek elkészítéséhez használunk. Ezek szerint az adatbázis elemeinek száma  $39 \times 39$  hangkapcsolódás, azaz 1521 elem. Ezen elemek gyakorlati létrehozásához el kell készíteni egy felolvasandó listát, amely tartalmazza az összes hangkapcsolódási

elemnek megfelelő szövegegységet. Az ilyen anyagot célszerű úgy megtervezni, hogy a bemondót ne befolyásolják az anyanyelvi kiejtési beidegződések. Ezért ezeket a listákat egyrésztől úgy állítják össze, hogy azok ne értelmes szavakat, hanem értelmetlen hangsorokat tartalmazzanak, másrésztől, hogy a hangsor- belseji elemekhez kivágandó elem ne az első szótag legyen (magyarban az első szótagot automatikusan hangsúlyozzuk), hanem például a második. A listák elemeinek összeállításánál figyelembe kell venni, hogy mely beszédhangok nem hatnak a másokra, illetve, hogy melyek hatnak és hogyan (Magdics 1965; Olaszy 1985). Célszerű például kihasználni, hogy a [k] hang nemigen befolyásolja az őt megelőző és az őt követő magánhangzó formánsszerkezetét (nincs átmeneti fázis a magánhangzóban, hanem a [k] illeszkedik a magánhangzóhoz). Így a magánhangzó formánsai a vágási pontokban viszonylag állandó értékűeknek várhatók. Ezzel a megoldással csökkenthetők a formánskülönbségek, amelyek az adott magánhangzó későbbi összeillesztésénél adódhatnak. A listában továbbá szerepeltetni kell az adott logatom mellett, hogy mely hangkapcsolati elem kerül kivágásra az elem felhasználásával. Célszerű azt is jelölni (a későbbi feldolgozás megkönnyítésére), hogy mely kivágási szabályt kell alkalmazni az adott elem elkészítésénél (lásd később). A 4. táblázatban példát mutatunk be a diádós adatbázis nyelvi anyagát tartalmazó listára.

4. táblázat: Nyelvi anyag diádós hangadatbázis készítéséhez

Sorszám	A felolvasandó szövegelem	A kivágandó Hangelem	A kivágás Szabálya
1.	# baka	# b	2
2.	abáka	bá	2,1
3.	aboka	bo, ka#	2, 1 és 2,1

A szabályok oszlopban a kivágási szabály (lásd később) sorszámát tüntettük fel a kivágandó hangok sorrendjében. Tehát az 1. elem kivágásánál a [b] hang zöngéjét kell kivágni a zár-felpattanásig a 2. szabály szerint, a 2. elemnél a [b] hangot a zár-felpattanás előtt, az [a:] hangot



pedig a közepénél kell elvágni a 2., illetve az 1. szabály alkalmazásával.

Hogyan építkezünk a diádós adatbázis-elemekkel? Mivel minden elem két félhangot tartalmaz, abszolút kezdő pozícióba mindig #-al kezdődő elem kerül, utána következnek a jel nélküli elemek, majd a hangsort olyan elemmel kell lezárni, amelyik # jellel végződik.

**Példa:**

Eredeti szöveg: *Jó napot kívánok.*

A diád elemek: #j-jó-ón-na-ap-po-ot-tk-ki-iv-vá-án-no-ok-k#

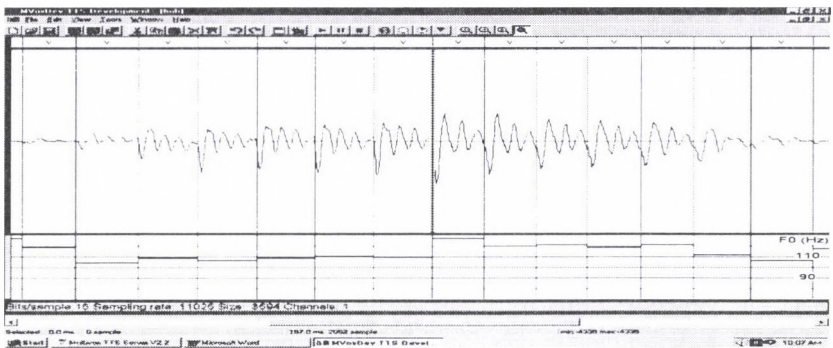
A példából két fontos kiejtési jellemzőt is kiemelünk. Az egyik, hogy látható, hogy a szavak határán nem #-kal jelölt elemek vannak, tehát a szavakat nem választjuk el egymástól hang szinten, hanem folyamatos beszédjelet hozunk létre. A másik pedig az, hogy figyelembe vettük a hosszan írjuk, röviden ejtjük szabályt a *kívánok* szóban, vagyis a *kí* szótagot a [k] és a rövid [i] kapcsolatával állítjuk elő.

### 2.2.2 A kevert szerkezetű adatbázis elemeinek tervezése

A kevert szerkezetű adatbázis felépítése bonyolultabb, mint a diádósé, mivel itt CVC típusú elemeket és diádós elemeket is tárolunk. Ennek a megoldásnak az az előnye, hogy jobb hangminőséget lehet vele elérni, mint a diádós adatbázissal, ugyanakkor ennek az az ára, hogy több elemet kell tárolni, illetve kezelni, szervezni a beszédépítés során. A CVC típusú elemek alkalmazását az a megfontolás indokolja, hogy a beszédben a magánhangzók hordozzák a leghangzósabb akusztikai információt. Törekedni kell tehát arra, hogy a magánhangzók akusztikai szerkezete minél korrektebbül legyen megvalósítva a beszédépítés során. Ezt a CVC szerkezet biztosítja.

A diádós szerkezetű adatbázisban a magánhangzókat ketté vágjuk és az utólagos összefűzésnél két fél magánhangzó találkozik egymással, aminek az is a következménye, hogy ezen magánhangzók formánsai nem követik azt a természetes vonalat, amit például két mássalhangzó közötti artikuláció megkívánna. A diádok kivágási módjának meghatározásánál (a hang felénél vágunk) tulajdonképpen azt a nem teljesen korrekt elvet követik, hogy egy C1-V-C2 kapcsolatban lévő

magánhangzónak három belső szakasza van, az átmeneti rész a C1-ből a V-be, a V tiszta fázisa és a V-C2 átmeneti rész. Ebből a felfogásból az következik, mintha a magánhangzó közepén egy rövid szakaszon változatlan, csak a magánhangzóra jellemző frekvenciaszerkezet lenne (ezt sugallja a tiszta fázis elnevezés). A gyakorlatban azonban ez a legritkább esetben fordul elő, döntően a hosszú magánhangzók ejtésekor. A rövid magánhangzóknál, mivel az artikuláció folyamatosan változik a C1 és C2 között, a spektrális szerkezetnek nincs állandó szerkezetű szakasza. Azt lehet itt mondani, hogy a magánhangzó minden periódusában a hullámforma más, nincsenek egyforma periódusok. Tehát bárhol vágjuk el a magánhangzót, a későbbi újraépítésnél a kapcsolódási pontoknál a periódusok belső szerkezete nem fog folyamatos változást mutatni. Ugyanez vonatkozik az alapprofrekvencia kérdésére is (3. ábra). Annak nagyon kicsi a valószínűsége, hogy a két összekapcsolt félmagánhangzóban az alapprofrekvencia értékek ugyanazok (ugyanis nincs olyan bemozdó, aki egy sok órás elemfelolvasásnál ugyanazon a Hz-értéken tudja tartani az alaphangját).

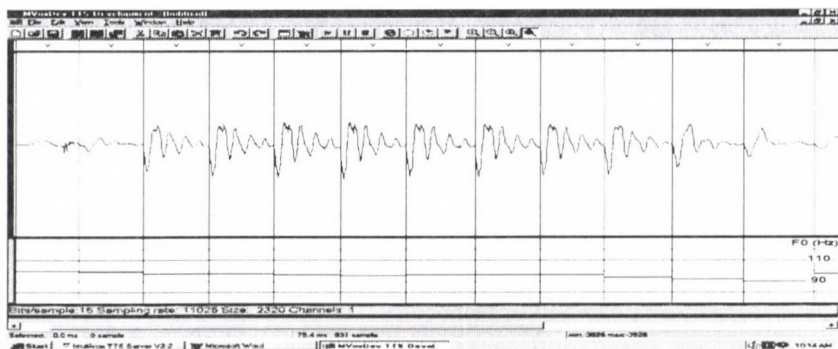


3. ábra

A *bob* hangsor [o] hangjának előállítására diádus adatbázis *bo* és *ob* elemével. Az [o] hang rezgése és alapprofrekvencia-görbéje

A fenti két torzítás a CVC adatbázis elemek bevezetésével kiküszöbölhető és biztosítható egyrészt, hogy a magánhangzó termé-

szetes formánszerkezete megmaradjon az egész hang folyamán, másrészről, hogy a hangon belül sem az amplitúdó menetben sem pedig alaphangfrekvenciában ne legyen törés (4. ábra). A 3. ábrán láthatjuk, hogy a két félmagánhangzó kapcsolódási pontján (a vastag vonal) a kapcsolódó periódusok szerkezete más. Továbbá látható, hogy a magánhangzó második fele intenzívebb periódussal indul (a 4,11 jelzés), mint amilyennel az első fele befejeződött. Az alaphangfrekvencia görbén látható, hogy a magánhangzó második felében az  $F_0$  értéke kb. 10 Hz-cel magasabb, mint az első felében.



4. ábra.

A *bob* hangsor [o] hangjának előállításához a *bob* triados elemmel. Az [o] hang rezgésképe és alaphangfrekvencia görbéje

A 4. ábrán láthatjuk, hogy ez az [o] hang mind rezgésképeiben, mind amplitúdó menetében, mind pedig alaphfrekvencia-szerkezetében megfelel a természetes beszédbeli szerkezeteknek.

### 3. A hangfelvétel elkészítése

A hangfelvétel a legérzékenyebb pontja az adatbázis-készítés sok-elemű műveletsorának. Gondoskodni kell arról, hogy a bemondó a hangerejét és hangmagasságát az egész felolvasás alatt lehetőleg azonos szinten tartsa. A fáradás ugyanis hangerő-csökkentésben és hang-



magasság-növekedésben jelentkeznek. A hangfelvételt célszerű maximum egy napra tervezni, ugyanis az ember hangja, hangszínezete naponta is változik.

#### **4. A diádós és triádós hangadatbázis elemeinek elkészítése**

A hangadatbázis elemeinek elkészítése azt a műveletsort foglalja magában, amelynek eredményeképpen előáll a hangsorépítéshez felhasználható hullámformák csoportja. Az elkészítés speciális szoftvereket igényel. Ez azt jelenti, hogy ilyen adatbázisok elkészítéséhez nem használhatók a hangkártyákhoz kapható hangeditáló szoftverek.

Ahhoz, hogy egy-egy hangadatbázis elem elkészüljön a következő műveleteket kell elvégezni speciális számítógépes programmal:

- a hanghatárokat be kell jelölni az eredetileg rögzített hangmintán,
- a zöngés hangokat el kell látni periódus jelzésekkel (markerek),
- a zöngétlen hangokba is markereket kell tenni, mégpedig a zöngés hangokból következő időszakazonként,
- az egyes markereket el kell látni zöngés (V), illetve zöngétlen (U) jelöléssel,
- ki kell vágni az adatbázis elemet az eredetileg felvett hangmintából adott szabály alapján és el kell helyezni az adatbázisban.

A fenti műveletek nem végezhetők el emberi beavatkozás nélkül, mivel a beszédjel egyrésztől nem szabályos rezgések sorozatából áll, másrésztől a beszédhangok határainak kijelölését sem lehet egyértelműen algoritmizálni. A markereket és a hanghatárokat pontosan kell elhelyezni, mivel a későbbi prozódiai feldolgozás (hangidőtartamok változtatása, a dallammenetek ráültetése a hangsorra, az amplitúdó-viszonyok megváltoztatása) ezeket az adatokat használja.

#### **A hanghatárok bejelölése**

Az adatbázis elkészítéséhez meghatározott szerkezetű hangsorokat olvastattunk fel a bemondóval. Ezt felhasználjuk a hanghatárok automatikus bejelölésénél. Készíthető olyan célszoftver, amelyik akár diádós, akár triádós elemekhez felvett hangsorokban az energia viszonyok alakulása alapján meghatározza a hanghatárok többségét, és el is

helyezi azokat a hangsorban. Vannak azonban olyan hangkapcsolatok, amelyekben az energia viszonyok nem változnak jellegzetesen a hanghatárokon. Ide sorolható a legtöbb CC kapcsolat, valamint a VV elemek, és a [j] hanggal képzett CV, VC kapcsolatok. Ezeknél a hangkapcsolatoknál kézi módszerrel audiovizuális célszoftver felhasználásával kell a hanghatárokat bejelölni.

### **A periódusjelzések elhelyezése a hangsorban**

A periódusjelzés célja az, hogy megjelöljük mindegyik zöngés hangperiódus kezdetét. Ez az alapja a későbbi hangnyújtásnak, hangrövidítésnek, valamint az alapfrekvencia változtatásának. Periódusjelzéseket elméletileg csak a zöngés hangokban lehet meghatározni, azonban a további feldolgozás egyszerűsítése céljából a zöngétlen hangokban is – noha ezekben nincsenek periódusok – célszerű ugyanolyan osztással markereket elhelyezni, mint amilyen osztással a zöngés hangokban tettük. (Ez például azért is fontos, mert a későbbiekben a hangidőtartam módosításokat a markerek felhasználásával végezzük el.) A periódusjelzések elhelyezése után tehát az egész hangsorban közel egyenletes osztással markerek találhatók. Célszerű a hanghatár jelzést és a periódusjelzéseket összehangolni, ami annyit jelent, hogy a hanghatárt mindig periódus jelzésre tesszük. Ebből következik, hogy minden beszédhang belseje egész számú kvázi egyforma távolságú markerrel lesz felcímkézve.

### **A zöngés-zöngétlen állapot jelzése**

A hanghatárokkal és markerekkel felcímkézett hangsort el kell látni zöngés-zöngétlen állapotjelzővel is. Ez azért fontos, mert az alapfrekvencia módosításokat zöngétlen hangokon nem kell végrehajtani. A fenti állapotjelzők meghatározása szintén jelfeldolgozó célszoftverrel automatizálható nagy részben. Az ilyen szoftverek számára a döntési nehézséget a kis energiájú hangrészek (például zöngés zárhang zárszakasza) és a kevert gerjesztésű hangok jelentik. Ebből az következik, hogy a feltehetően hibás döntésű elemekben kézzel kell a megfelelő állapotjelzőket beállítani.

### **Az elemek kivágása és elhelyezése az adatbázisban**

Az elemek kivágása nagy részben szintén automatizálható, mivel azt jól meghatározható szabályok szerint kell elvégezni. Diádus rendszerű elemekhez általános szabály lehet az, hogy a magánhangzókat a közepük táján kell elvágni, mindig periódusjelnél, ami biztosítja azt, hogy a hang időfüggvényének menete nem fog megtörni a későbbi hangsorépítés során. Ez azt jelenti, hogy például egy zöngés hangelemben a vágási ponton az időfüggvény a nulla vonalat alulról közelíti, majd a hozzá kapcsolt másik zöngés elemben az időfüggvény a nulla vonalról felfelé indul, tehát folyamatos lesz az átmenet az elemhatáron. A zárhangokat, és a zárrés hangokat a zárfelpattanás előtti jelzésnél kell elvágni, tehát nem a hang közepén. Ez azért jó, mert így ezen hangoknál a hang nyújtása, rövidítése egyszerűbb.

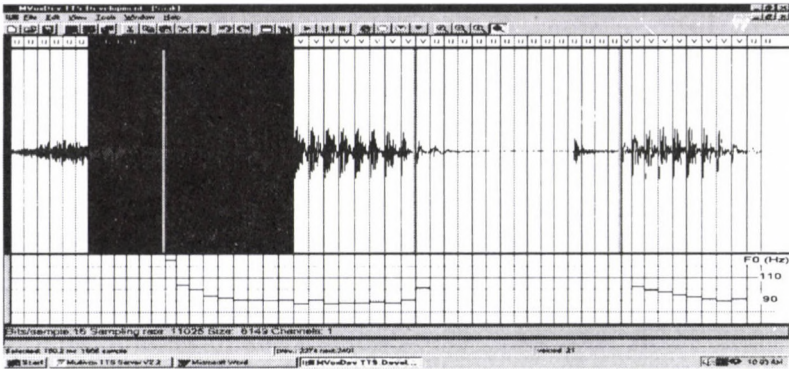
Diádus beszédépítéshez háromféle diádot kell létrehozni: hangsor kezdőt, hangsor belsejit és hangsor zárót. A hangsorkezdő elemek a #C, #V, a hangsorzárók a C#, V#. A hangsorbelseji elemek a CV, VC, CC, VV. Ennek megfelelően a következő szabályok alapján kell kivágni a diádus adatbázis-elemeket az eredetileg felvett beszédelemekből.

#### **1. Szabály: a vágási pont a hang közepén van (5. ábra)**

- a #V, #C, C# és V# elemekben ahol a C réshang és az [m, n, l] hang
- a CV, VC elemekben ahol a C réshang és az [m, n, l] hang,
- a VV elemekben
- a CC elemekben ahol a C réshang és az [m, n, l] hang.

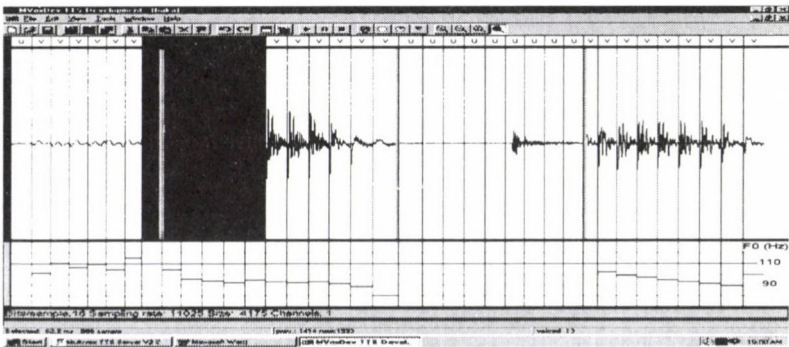
Az [r] hang egyéni elbírálást kíván, mindig az adott bemondó hangképzésétől függ, hogy mely ponton kell vágni.





5. ábra.

A *sá* diád kivágása a *sáka* hangsorból (a sötétben kijelölt rész)



6. ábra

A *ba* diád kivágása a *baka* hangsorból (a sötétben kijelölt rész)

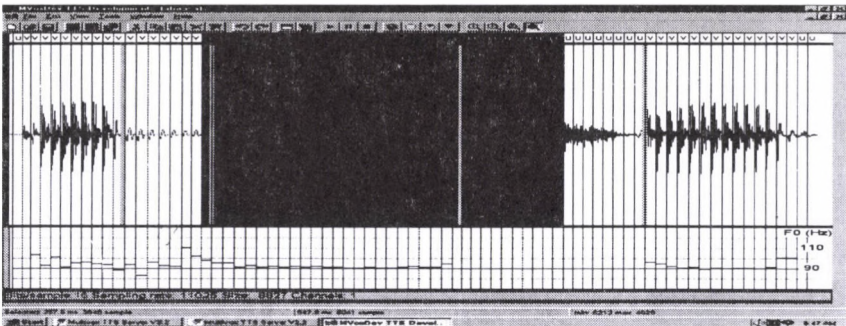
**2. Szabály:** a vágás a zárfelpattanás előtti ponton van (6. ábra)

- a #C, C# elemekben, ahol C zárhang, illetve zár-rés hang,
- a CV, VC elemekben ahol a C zárhang, illetve zár-rés hang,
- a CC elemekben ahol a C zárhang, illetve zár-rés hang

A triádös szerkezetű elemek készítéséhez más szabályok szerint kell

eljárni. Triádos elemekként kezeljük a hangsorkezdő #VC, a hangsor-belseji CVC és a hangsorzáró CV# hangkapcsolatokat.

Itt általános szabály lehet az, hogy a magánhangzókban nincs vá-gási pont, csak a mássalhangzókban és ezekre ugyanazok a szabályok vonatkoznak, mint amilyeneket a diádós elemek mássalhangzóira megadtunk (7. ábra).



7. ábra

A *bás* triád kivágása az *abása* hangsorból (a sötétén kijelölt rész)

A 7. ábrából látható, hogy a [b] hangban a vágási pont a zár-felpattanás előtt van, az [ʃ] hangban pedig a hang közepén.

## 5. Akusztikai össze-síszolás

A fenti szabályok alapján elkészített adatbázis elemeket még nem lehet felhasználni közvetlenül beszédépítéshez, mivel azok nincsenek egymáshoz illesztve sem amplitúdóban, sem hangidőtartamban. Ez az állapot abból adódik, hogy az elemeket emberi bemondásból származtatjuk és a bemondó hangereje, beszédtempója általában változik a hosszú elemlista felolvasása során, még akkor is, ha professzionális bemondót alkalmazunk. Így, ha összekapcsolnánk ezeket a „nyers” elemeket a beszédépítés során, akkor torz és lötyögős hangzást kapnánk, olyant, mint amilyent a 3. ábrán mutattunk be. Az akusztikai

csiszolás lényege, hogy az egyes adatbázis elemeket összekapcsoljuk más elemekkel (a legideálisabb, ha mindegyiket mindegyikkel), és az amplitúdókat egymáshoz igazítjuk, továbbá a hangidőtartamokat is beállítjuk egy ún. specifikus időtartamra. Erre azért van szükség, mert a beszédben a hangok egymáshoz való viszonyának helyes aránya a mérvadó. Ha ezek az arányok jók, akkor folyamatosnak és simának, egyszóval természetesnek halljuk a beszédet, ha viszont nem helyesek, akkor zötyögősnek, szaggatottnak, lüktetőnek. Jelenleg ez a legmunkaigényesebb fázisa az adatbázis készítésnek, mivel ezt nemigen lehet automatizálni. Nem is beszélve arról, hogy számos elemben kell periódusokat is javítani, mert például a bemondó hangja éppen kissé reszelőssé vált, túl levegős lett stb. Egyszóval az adatbázis csiszolása során meghallgatjuk az összes elemet (környezetbe ágyazva) és így döntjük el, hogy mely pontokon kell javítani és mit. Gyakorlati tapasztalat az, hogy 15-20 oldal szöveg szintetizálásával a leggyakoribb hangkapcsolatok hangzását meg lehet vizsgálni és el lehet végezni a csiszolást. Statisztikai gyűjtéssel ezeket dokumentálni lehet, és utána a fennmaradó, ritkábban előforduló elemeket külön kell vizsgálni és összehcsiszolni.

### Irodalom

Hirokawa, T.–Itoh, K.–Sato, H.: High quality speech synthesis based on wavelet compilation of phoneme segments. Proc. of ICSLP 1. Banff, Alberta, Canada 1992, 567-570.

Magdics Klára: A magyar beszédhangok akusztikai szerkezete. Akadémiai Kiadó. Budapest 1965.

Németh, G.–Ferenczy, T.–Olaszy, G.–Gáspár, Z.: A flexible client-server model for multilingual CTS/TTS development. Proceedings of Eurospeech '97. Vol. 5. Rhodes 1997, 693-697.

Olaszy Gábor: A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise. NyttudÉrt. 121. Budapest 1985.

Olaszy Gábor: Elektronikus beszédelőállítás. Műszaki Könyvkiadó. Budapest 1989.

Ez a kutatás az OTKA T 03292 támogatásával készült.



# **„JÓ LENNE EGY SZAKÉRTŐTŐL MEGTUDNI, HOGY MELYIK AZ IGAZI ANYANYELV” TÖBBNYELVŰSÉG ÉS ÖNAZONOSÍTÁS**

**Navracsics Judit**

**Veszprémi Egyetem, Alkalmazott Nyelvészeti Tanszék**

## **Egyéni többnyelvűség és identitás**

Az anyanyelv és az önazonosítás meghatározásakor azok a két- vagy többnyelvűek vannak a legnehezebb helyzetben, akik nem tartoznak nemzeti kisebbséghez, hanem valamilyen családi vagy munkájukkal kapcsolatos ok miatt váltak két- vagy többnyelvűvé. A Veszprémi Egyetemen évek óta folynak egyéni két- és többnyelvűségi kutatások. Kísérleti személyeim különböző okokból kerültek Magyarországra, néhány dolog azonban közös bennük: megtanultak magyarul, Magyarországon élnek, tanulnak vagy dolgoznak, családi és/vagy baráti kapcsolatot tartanak fenn. A címben idézett óhaj egy magyar–angol 26 éves férfitől származik. Ez késztetett arra, hogy alaposabban megvizsgáljam, hogy miként is történik az önazonosítás az egyéni többnyelvűeknél, melyek azok a meghatározó tényezők, amelyek eldöntetik az egyénnel, milyen nemzetiségűnek, milyen anyanyelvűnek tartja magát.

1. Egy angol–perzsa–magyar háromnyelvű testvérpár nyelvi fejlődését kísértem figyelemmel négy éven keresztül. Az anya perzsa–angol kétnyelvű, az apa angol egynyelvű. A családnyelv az angol, a gyermekek közös nyelve a magyar, miután magyar bölcsődébe, majd óvodába, most pedig már iskolába járnak. Többnyelvűségük összetétele erősen megváltozott a négy év alatt. Legtöbbet a magyart használják és a magyar nyelvi kompetenciájuk a legjobb. A szülők a Kanadába való hazalátogatást tervezik, hogy egy kicsit javítsanak a gyermekek angol tudásán. A gyermekek perzsa nyelven csak az édesanyjukkal tudnak kommunikálni. Ha valami bajuk van, perzsául szólnak anyjukhoz, és az anya számára is egyértelmű, hogy ha dicsérni

vagy éppen megróni akarja a gyermekeket, akkor a perzsa nyelvet használja. Mi tehát ezeknek a gyermekeknek az anyanyelve vagy első nyelve? Az angol, amit legelőször tanultak meg, a perzsa, amelyen érzelmi életüket élik vagy pedig a magyar, amelyet legtöbbször használják kortársaikkal, a befogadó társadalommal, és amelyen a társas érintkezésüket fenntartják?

2. Több kamaszkorú többnyelvű kísérleti személyem szülei azonos, de nem magyar anyanyelvűek (például vietnámi, kínai, német, orosz, angol). A családi nyelvhasználatra magyarországi érkezésükkor a szülők anyanyelve volt a jellemző, azonban amiatt, hogy a gyermekek magyar közösségbe járnak, a család nyelvhasználata is változik. A gyermekek magyar nyelvi kompetenciájuk fejlődésével egyre inkább otthon is kikövetelik a magyar nyelvhasználatot. Ugyanez vonatkozik a kultúrára is. Amíg kezdetben kizárólag a szülők kultúráját tartották magukénak, mára már a kultúra területén is változások állnak be. Sajátos kultúrát alakítanak ki, amely két- vagy többféle kultúrát ötvöz.

3. Bár minden korosztályt ösztönösen foglalkoztat az önazonosítás kérdése, mégis talán a felnőttek azok, akik elé a társadalom olyan tükröt tart, hogy adott esetben meg is kell fogalmazniuk hovatartozásukat. Felnőtt korú kísérleti személyeimet két csoportba lehet osztani:

3.1. A vizsgált többnyelvűek közül az egyik csoportba azok a – többnyire – nők tartoznak, akik más országból érkeztek Magyarországra 20-as éveik elején feleségként. Az, hogy hamar munkába álltak, rákényszerítette őket arra, hogy minél előbb megtanuljanak magyarul. A tanulási folyamatot az is gyorsította, hogy gyermekeikkel csak addig beszéltek az anyanyelvükön, amíg azok bölcsődébe vagy óvodába nem kezdtek járni. Amint megkezdődött a gyermekek szocializációja a magyar társadalomba, az anyák anyanyelve egyre hátrányosabb helyzetbe került. Nem tűnt el teljesen, de korlátozott szituációkban tudták már csak használni. Hogy mennyire volt erős a motiváció arra, hogy ezek a személyek megtanulják a magyar nyelvet, álljon itt néhány vallomás:

„... úgy éreztem magam, mint aki süketnéma ...” (M., 43 éves német első nyelvű, német–magyar kétnyelvű)<sup>1</sup>

„... úgy éreztem magam, mint egy üvegburában. Körülöttem az emberek tátogtak, és én semmit nem értettem.” (T., 50 éves lengyel első nyelvű, lengyel–orosz–magyar háromnyelvű)

Az önazonosítással kapcsolatban azonban a következőket mondják:

3.1.1. M.: „Ha németül beszélek, akkor úgy érzem, én vagyok, ha magyarul, akkor más. Olyan, mint hogyha egy kabátot felveszek, akkor nem olyan vagyok, mint természetesen. Az idegen.”

Tehát, annak ellenére, hogy több mint 20 éve Magyarországon él, és többnyire magyarul beszél, idegennek érzi magát itt, németnek vallja magát.

3.1.2. T.: „3 nyelv él bennem párhuzamosan, egymástól függetlenül. Bármelyik nyelvre átléphetek akadály nélkül.”

Ugyanakkor vannak olyan nyelvi tartományok, amelyeket csak az egyik nyelven ismer. Magyarországon tanult meg autót vezetni, itt születtek a gyermekei, az autóval, gyermekbetegségekkel kapcsolatos lexikát csak magyarul tudja. Számára az országtól függ, hogy melyik nyelv a természetes: Lengyelországban a lengyel, Oroszországban az orosz, Magyarországon a magyar. Ezek után azt hihetnénk, identitását is a helyzetnek megfelelően határozza meg. Azonban a kérdésre, hogy ő maga minek tartja magát: lengyelnek vagy magyarnak, rövid gondolkodás után azt válaszolta, hogy magyarnak. Annyi év eltelt már, amióta Magyarországon él (29 év), hogy „teljesen beléivódott” a magyar lét. Állítása szerint szinte mindent magyarul csinál: magyarul gondolkodik, magyarul számol, magyarul álmodik. Egy dolgot



azonban csak lengyelül tud: imádkozni. Magyarul is ismeri az imákat, de imádkozni csak lengyelül szokott.

3.1.3. K. (63) görög–magyar kétnyelvű személy jól érzi magát Magyarországon, kettős állampolgárságú, a társadalomba beilleszkedett, saját bevallása szerint: „tökéletesen beszélek magyarul is és görögül is” (holott éppen az ő nyelvtudása a leggyengébb az adatközlőim között!), „nem érezzük, hogy idegenek vagyunk.” A társadalmi beilleszkedés, az otthon érzete elhomályosítja a nyelvi problémákat, alanyom nem is foglalkozik a nyelvi pontatlanságaival. Mindkét kultúrában otthon van, mindkét nyelven kommunikációképes, a kétnyelvű, kétkultúrájú egyén grosjeani (1992) kritériumainak megfelel.

3.2. Adatközlőim másik csoportját olyan fiatalok alkotják, akiknek magyar anyanyelvű szülei Kanadába, az USA-ba és Angliába emigráltak. Mindegyik személy arról számolt be, hogy az otthon nyelve a magyar volt mindaddig, amíg iskoláztatásuk el nem kezdődött.

3.2.1. P. (28 éves magyar–amerikai) vallomása a következő:

P.: „Ötéves voltam, akkor mentem iskolába, s tanultam angolt és akkor ööö nagyon akartam elfelejteni a magyart.

KV: Igen, de miért?

P.: Hát mért? *Cos* senki más beszélt magyart, csak angolt és ööö valahogy én akartam lenni mint más, igen mert barátomnak is és akkor mikor hazamentünk ööö nem akartam magyart beszélni és csak angolt és kiderült hogy csak angolt, *I mean* magyart beszéltünk ... „.

Tehát társai, barátai előtt szégyellte, hogy nem olyan, mint a többiek. Ettől kezdve a családi nyelv kevert volt, így P. nagyon sokat vesztett el magyar tudásából. Testvére még jobban ellenállt a szülői törekvéseknek:

P.: „Ő rosszabb volt, mint én, ő nagyon nem akart. Ő más, mint én ... nem magyar, nem lett magyar, nem is kell neki most, tudod?”

A szülők nem adták fel, szerették volna magyar iskolába beíratni a fiúkat:

P.: „.... egyszer akart a szüleim, hogy menjünk egy magyar iskolába. És mentünk első nap, és én és az öcsém bemászkáltunk egy szobába és ööö ott volt magyarok és táncoltak ööö régi ruhákban, magyar ruhák, tudod, nagyon régi és ilyen magyar tánc és én és az öcsém: ez nem lesz. Nem – azt mondtam apámnak és anyámnak: azt hiszed, hogy én ebbe az iskolába megyek? ... Komolyan mondtam neki, hogy én nem megyek ebbe egy percig. So, visszamegyünk a másik régi iskolámba és hagyjuk ezt, tudod. Ilyen magyar, aki táncol, én nem akarok semmit róla, tudod?”

Itt a magyar kultúra egy szeletétől való idegenkedésnek vagyunk tanúi. Meg kell azonban jegyezni, hogy az emigráns csoportokban fenntartott kulturális jellegű tevékenységek valóban kicsit mesterkéltek és lehet, hogy egy magyar egynyelvű, egykultúrájú gyermek is eljedt volna az adott iskolától.

P. sorsa azonban úgy alakult, hogy egy érzelmi válság idején a legjobb megoldásnak azt látta, hogy eljöjjön Magyarországra. 27 évesen újra kezdte az életét szülei hazájában (akik most is az USA-ban élnek). Nagyon szeret itt lenni, újra tanulja a nyelvet, magyar barátai vannak, szereti a magyar tájat, Budapestet:

P.: „Tudod? *Culture* van. Mondjuk én a New York City-ben is laktam és ott is nagyon volt mindenhol levegő, tudod, érezted, tudod? És itt is valahogy ilyesmi van. Nem olyan sokat, olyan sok, mint a New York nekem, csak más. Itt inkább a *culture* és *arts* tetszik nekem, épület, ... mikor megyek a Dunába és nézek a Budán, akkor nagyon gyönyörű nekem...”

P. nyilatkozatai alapján kétnyelvű és kétkultúrájú embernek tekinthető. Ismeri a két kultúra lényegét, s a kettőből egy sajátos egységet

formál. Nem tekinthető sem magyarnak, sem amerikaiak, mégis egységes egész.

3.2.2. S. (26 éves magyar–angol) szülei Angliába vándoroltak ki még fiuk születése előtt. A családnyelv szintén a magyar volt.

S.: „... mikor először jártam az óvodába, akkor én nem is nagyon tudtam angolul. Ott kellett nekem megtanulni azért, mert otthon általában vagy akkor még csak magyarul beszéltünk. Szóval, én én angolul csak a három vagy négy éves voltam, amikor valóban nekiálltam.”

Az iskoláskor idején már egyre többször beszélt S. otthon is angolul, hiszen már nem mindent tudott elmondani magyarul szüleinek arról, hogy mi történt az iskolában. Az identitással a következő problémái voltak:

S.: „Hát, az az igazság, hogy én nagyon másnak éreztem magam. Az tény, hogy az iskola nagyon nemzetközi jellegű volt, és ahogy meséltem, sok volt a lengyel, pakisztáni, amerikai. Szóval nagyon vegyes társaság, de én soha nem éreztem soha nem éreztem magam, hogy mondjuk echte britnek. Tudtam, hogy más a nevem, tudok egy furcsa nyelvet, aki senki más abba az iskolába nem tudott. És a szüleimtől, szóval otthon magyarul beszéltünk, magyar ételeket ettünk, szóval magyarnak éreztem magam, de gondoltam, hogy milyen a Magyarország?”

Amikor ellátogatott Magyarországra, éreznie kellett, hogy itt is más, hiszen szülei magyar 56-os vagy még korábbi nyelvezetét használta a magyar egynyelvű unokatestvérek nagy multságára.

Angliában élni, és munkát kapni – származása miatt – nem volt „paradicsom”. Ő volt az egyetlen a szűk baráti körből, aki kapott állást az egyetem után Londonban.



A vasfüggöny megszűnése után a család hazatért. S. Budapesten, a szülők Sopronban élnek. Egy kirándulásra így emlékszik vissza:

S.: „Azt hiszem, szeptember volt, sütött a nap és én meg vezette és muszáj volt megállnom csak kiszállni és körülnézni és mondtam az apámnak, hogy ez egy ilyen, ez olyan gyönyörű, mint bármelyik spanyol, francia vidék, ami, ahol én jártam és kiderült, annak idején én nem tudtam, de ahol megálltunk, az a mádi, az Mádhoz közel és a családnak az egyik ágazata ott lak, ott voltak elég nagy szőlő, hogy mondjam, telep. Szóval, ez volt, nem csak, lehet úgy is magyarázni, hogy a családi háttér vagy valami vonzott, ami nem pusztán a környék.”

Ezek után hihetnénk, hogy érzelmileg Magyarországhoz jobban kötődik, ám egy elszólás alkalmával Angliát nevezi hazájának (igaz, azonnal javítja is a nyelvbtlást: „hazam” (egyek) helyett „Angliába megyek”-re):

S.: „... mindig szoktam venni, ha hazam (helyesbít), Angliába megyek.”

A konkrét kérdésre, hogy magyarnak vagy angolnak vallja-e magát, a következő választ adta:

S: „Se magyar, se angol és ez a nehéz benne. Angliába nagyon jól, szóval nincs ott problémám gyakorlati dolgokban, meg tudom értetni magamat, jól tudom a nyelvet, több mint húsz évet éltem ott ... de mégis van mellette egy plusz, amitől másnak érzem magam. Én bennem nem volt az az óriási vágy, hogy a brit legyek, hogy teljesen annak, az legyen belőlem, hogy egy, egy angol. Szóval, ez az egyik. A másik, hogy itt vagyok most Magyarországon, magyar útlevelem van, magyar fizetést kapok, de mégis nem lehet, nem tudok úgy viselkedni, hogy valaki ne vesz..., ne vesz..., ne tudja, hogy nem, mert világos, hogy nem az átlag magyarról van

szó, nem 100 %-os az a magyarság vagy az a magyar tudat, vagy a viselkedési módom.”

Az egynyelvű szemléletre ad kritikát S., amikor megfogalmazza, hogy az itteni magyarok másképp viszonyulnak hozzá, mint egynyelvű társaihoz. Az a tudat, hogy máshonnan jött, kettős állampolgárságú, nemcsak benne épít ki gátat, hanem a magyarokban is. „Az ember nem tud integrálódni addig, ameddig a itt is, hogy mondjam akceptálják az embert.”

Ebben az esetben is a kétnyelvű, kétkultúrájú ember példáját láthattuk.

### **Csoportos többnyelvűség és identitás**

Az egynyelvű államokban ritkán merül fel a kérdés, hogy vajon az állam területén élőknek mi is az anyanyelve vagy éppen milyen nemzetiségűek. Természetesnek veszik az egy nemzet – egy nyelv – egy kultúra triászát. Ezen a tényen azonban minél előbb változtatni kell. Legelőször is fel kell hagyni azzal a meglehetősen felületes és általánosan elfogadott euro-atlanti nézettel, miszerint az egynyelvű államok számszerűen felülmúlják a többnyelvűeket és az egynyelvűség a norma. Ha kitekintünk a világba, láthatjuk, hogy a két-, illetve többnyelvűség mindenhol jelen van, és a két-, többnyelvű országok száma nagyobb, mint az egynyelvűeké. Nemzetközi jogi dokumentumok (például a Hágai, vagy az Osloi Ajánlások) biztosítják a nemzeti, nyelvi kisebbségeknek arra vonatkozó jogát, hogy saját nyelvüket használhassák, saját kultúrájukat ápolhassák. Ezek a dokumentumok azonban nem minden országban jutnak el az érintettekhez (házánkban is egy kulturális folyóirat, az Új Horizont jelentette meg mellékletként a Hágai Ajánlásokat). Számos országban tapasztalhatjuk a kisebbség egyre nagyobb mértékű asszimilációját a többségi társadalomba. Az asszimiláció okai különbözőek lehetnek, kezdve a politikaitól egészen a magán jellegűig.

A kisebbségek számára a történelem sokszor „megkönnyítette” az identitással kapcsolatos kérdésre a választ. Ismeretes, hogy az anya-

nyelv és a nemzetiség bevallása sokszor politikai következményekkel járt. A következő táblázat a Magyarország területén élő nemzeti kisebbségek, azaz csoportos kétnyelvűek adatait tartalmazza. A számok változása is feltehetőleg vagy a diszkriminatív társadalmi tényezőket (még ha itt ezt részletesen nem is elemzem) vagy az egész kérdéskör vizsgálatának szaktudományi bizonytalan alapjait tükrözi.

1. táblázat: Az 1980 és az 1990 évi népszámlálás kisebbségi adatai\*

Kisebbségek	Anyanyelv szerint		Nemzetiség szerint	
	1980	1990	1980	1990
szlovák	16054	12745	9101	10459
román	10141	8730	8874	10740
horvát	20484	17577	13895	13570
szerb	3426	2953	2805	2905
szlovén, vend	3142	2627	1731	1930
német	31231	37511	11310	30824
cigány	27915	48072	6404 (sic!)	142683
örmény	"-"	37	egyéb: 16369	egyéb: 19640
görög	"-"	1640		
bolgár	"-"	1370		
lengyel	"-"	3788		
ukrán, ruszin	"-"	674		
összesen	112 393	137 724	70 489	213 111

\*Forrás: MAPSTAT Központi Statisztikai Hivatal szoftver, Budapest, 1992.

Látható, hogy az esetek túlnyomó többségében az anyanyelv szerinti meghatározás nagyobb számú mind az 1980-as, mind pedig az 1990-es népszámlálás alkalmával, mint a nemzetiség szerinti. Ugyanakkor azt a tendenciát is felfedezhetjük, hogy mind az anya-



nyelv szerinti, mind pedig a nemzetiség szerinti bevallás száma is csökkent a tíz év alatt. Ez alól kivételek a német és a cigány kisebbség adatai (erre még később visszatérek). Az adatok nem szolgáltatnak bizonyítékot az anyanyelv és a nemzetiség bevallásának összefonódására. Többen tartják magukat a kisebbségi nyelv hordozójának, mint a kisebbségi kultúra képviselőjének. Az általános konklúzió a fenti adatok alapján pedig csakis az, hogy az asszimiláció egyre növekszik.

### **Egynyelvűség, egy kultúrájúság**

Egynyelvű az az egyén, aki egynyelvű környezetben született, ott nőtt fel és a mindennapi életében egyetlen nyelvet használ, másra nincs szüksége. Anyanyelve meghatározása egyértelmű, az a nyelv, amit édesanyjától (és a közvetlen környezetétől) tanult meg. Kultúrája szoros kapcsolatban van nyelvével, egy kultúra szabályait tartja be, ahhoz tartozónak vallja magát, annak a kultúrának az ünnepeit üli meg, annak viselkedésszabályait és etikettjét tartja be.

### **Többyn nyelvűség, többkultúrájúság**

Többyn nyelvű az az egyén, aki többyn nyelvű közegben született, ott nőtt fel és a mindennapi életében két (vagy több) nyelvet használ az élet különböző területein, különböző szituációkban, különböző emberekkel (vö. Grosjean 1992). A két- vagy többyn nyelvű ugyanolyan értékű egyén mint az egynyelvű, egy nyelvileg kompetens kerek egész, akinek nyelvi kompetenciáját azonban nem lehet összehasonlítani az egynyelvűekével. Nyelvtudásuk minőségileg nem több, nem kevesebb: egyszerűen más. Egy sajátos összetétel, amely mindkét (vagy több) nyelvből tartalmaz elemeket. Grosjean a sportból vett hasonlattal él: a gátfutót sem lehet összehasonlítani a rövidtávfutóval sem a magasugróval, a gátfutás mégis az atlétikának önálló része! Ahogy a gátfutás is a természetes mozgás egyik fajtája, ugyanúgy a két- (vagy többyn) nyelvű fejlődés is természetes.

A kétnyelvűek anya- (első) nyelvét sok esetben nehéz meghatározni. Vajon mi az anyanyelve annak a kétnyelvűnek, aki kisebbségi te-

rületen született, de mivel az egyik szülője (az apa) a többségi nyelv beszélője, a családn nyelv is a többségi nyelv? Vagy mi az anyanyelve annak a kétnyelvűnek, akinek a szülei különböző nyelveket beszélnek és a társadalomban egy harmadik nyelv a domináns? Csak néhány példán keresztül is láthatjuk, hogy indokolt az „anyanyelv” kifejezés kiiktatása, sokkal helytállóbb az „első nyelv” megnevezés. Így már könnyebben dönthet a kétnyelvű a nyelveit illetően. Tudja, melyik az első nyelve vagy azért, mert azt tanulta meg először vagy pedig azért, mert azon jobban ki tudja magát fejezni, vagy azért, mert érzelmeit azon a nyelven fejezi ki. A nyelvek tehát különböző funkciót töltenek be és valószínű, hogy az anyanyelv – nem anyanyelv oppozíció nem is értelmezhető a két- és többnyelvűek számára: ez a dichotómia csak az egynyelvűek nézőpontjából vetődik fel.

A többnyelvű ember általában többkultúrájú is. Ez azt jelenti, hogy – szerencsés esetben – ötvözi a két vagy több kultúrát, amelynek része, és egy sajátos kulturális konfiguráció képviselője lesz. Fontos azonban, hogy a harmónia meglegyen, hogy az egyén semmiféle tudathasadásos állapotba ne kerüljön, azaz ugyanúgy, mint nyelvi állapota, kulturális állapota is egy kerek egész legyen, és ebben különbözzön a két vagy több egykultúrájúságtól.

### **Az egyén önazonosítása**

Ha az 1. és 2. pontban leírtakat vesszük alapul, látszatra nem nehéz meghatározni az egyén anyanyelvét, nemzetiségét, kultúráját. De vajon miért állítja olyan határozottan Petőfi – a szerb apa és szlovák származású anyja fia – versében: „Magyar vagyok.”? Ugyanakkor miért vallja magát Zrínyi Miklós horvátnak, miközben verseit magyarul írja? Melyek azok a pszichológiai tényezők, amelyek az egyénnel eldöntetik, milyen nyelvűnek, milyen nemzetiségűnek vallja magát?

### **Az identitás mint változó kategória**

A két-, illetve többnyelvű identitástudat feltehetően történelmileg változó kategória. Voltak idők, amikor például a vallás (lásd csángók)



vagy a kultúra presztízse (lásd Zrínyi és Petőfi idején a magyar kultúra egyenlő a nemesi kultúrával) határozták meg az egyén identitását.

Ma már inkább: a gazdasági hatalom (például az 1980–1990-es népszámlálások között a német kisebbség száma jelentősen növekedett az anyaországban beállt politikai és gazdasági változások hatására); az iskoláztatás; a nyelv; a földrajzi helyzet; a társadalmi szerep, társadalmi hozzáállás (a romák helyzetének javulása a 80–90-es években); a kultúrához való kötődés látja el az identitást meghatározó funkciót.

Mi határozza meg tehát az egyén számára az identitást? A nyelv? A család? A környezet? A kultúra?

Az ebben a tanulmányban elmondottakból arra a következtetésre lehet jutni, hogy egyik sem igazából és döntően, hanem mindegyik valamelyest hozzájárul az egyén teljes önképéhez. Hiszen hiába tud jól magyarul a német–magyar kétnyelvű hölgy, hiába magyar a családja, magyar a környezete, ő mégis, még 20 év után is németnek tartja magát. Ugyanakkor a lengyel–oroszmagyar háromnyelvű hasonló körülmények között magyarnak vallotta magát. Nem tartotta magát magyarnak az amerikai fiú, aki idegenkedett a magyar nyelvtől és a kissé avitt magyar kultúrától, most azonban vállalja magyarságát, kétnyelvűségét és kétkultúrájúságát. Hiába volt angol tudása tökéletes S.-nek, mégsem akart tökéletesen brit lenni. Büszke volt arra, hogy valamiben más, mint a többiek.

Egyben megegyeztek az adatközlők: mindegyik büszke kétnyelvűségére, hangsúlyozzák a kétnyelvűség előnyeit: aki igazi kétnyelvűként nő fel, annak sokkal könnyebb más nyelveket megtanulni. Sokkal több a kétnyelvűség előnye, mint hátránya. A hátrányt sajnos az egynyelvű társadalom okozza: amikor úgy néznek a kétnyelvűre, mint valami csodabogárra, valakire, aki teljesen más, mint a többi. Ez egyes esetekben serkentőleg hathat (lásd S. esetét), mert tetszik a másság, eltérő esetekben (lásd P. esetét) zavarja a fejlődésben levő egyéniséget, aminek eredménye vagy a származás teljes megtagadása (lásd P. testvére esetét) vagy pedig az ősök felkutatása, a kultúra iránti érzékenység megnyilvánulása (lásd P. esetét).



Zárszóul álljon itt S. gondolata, amely igazolja, hogy magukat a kétnyelvűeket is foglalkoztatja az anyanyelvük, nemzetiségük kérdése, meghatározása:

S.: „Számolni most már mind a két nyelven, szóval, ha valamit számolok, akkor néha angolul csinálom, néha magyarul. Elég gyakran beszélek az álom, szóval mikor alszom az álmaimban. És nagyon sokszor, azt hiszem 50%-a magyarul beszélek, szóval, jó lenne egy szakértőtől megtudni, hogy melyik az igazi anyanyelv. Lehet, hogy az én esetemben nem is lehet megmondani. Nem tudom.”

<sup>1</sup>Az interjúk archivált formában megtekinthetők a Veszprémi Egyetem Alkalmazott Nyelvészeti Tanszékén.

### Irodalom

Grosjean, F.: Another View of Bilingualism. Cognitive Processing in Bilinguals. Elsevier Science Publishers. Amsterdam 1992, 51-62.

Grosjean, F.: The bilingual and bicultural person in the hearing and in the deaf world. Sign Language Studies 77. 1992, 307-320.

Grosjean, F. (ed.): Living with Two Languages and Two Cultures. Paradis, I. Cultural and Language Diversity and the Deaf Experience. CUP Cambridge, New York 1996, 20-37.

A nemzeti kisebbségek oktatási jogai. Hágai Ajánlások és Értelmező megjegyzések. Új Horizont. XXVI./1. Melléklet. 1999, 1-12.

A nemzeti kisebbségek nyelvi jogai. Oslo-i Ajánlások és Értelmező megjegyzések. (megjelenés alatt)

# **HANGZÓ BESZÉD – ÍROTT BESZÉD: KOMMUNIKÁCIÓ**

**Lengyel Zsolt**

**Veszprémi Egyetem Alkalmazott Nyelvészeti Tanszék,  
JPTE Nyelvtudományi Tanszék**

E rövid írás – különböző területek empirikus adataira támaszkodva – amellett kíván érvelni, hogy az ember nyelvi képessége egy egészet alkot. Ez a kompakt nyelvi képesség különböző konkrét (hangzó, írott, taktilis stb.) manifestációt nyerhet. A különböző megjelenési formákat mind az egyén, mind a társadalom kiaknázza. Ily módon egyfelől integrált egészből beszélhetünk, miközben a nyelvi képesség különböző megjelenési formái meglehetősen nagy autonómiát élveznek. Mind az integrált, mind az autonóm jelleg további előnyöket biztosít. Az egyén szempontjából a különböző nyelvi formák, a nyelvi képesség különféle konkrét megjelenési formái a személyiség gazdagodását, kommunikációs kompenzációs technikák kidolgozását jelentik. Társadalmi szempontból a társadalmi jellegű közlések finomabb strukturálását, a humán kommunikáció nyelvi infrastruktúrájának bonyolultabbá válását, ezáltal a jóval bővebb információ-átadást teszik lehetővé.

## **Viszony, diszciplináris „fontosság”**

Az emberi kommunikáció – más komponensek mellett – magába foglalja a nyelv hangzó és írott megjelenését, megjelenítési formáját is. A kettő közötti különbségtétel ritkán tűnik rendszeresen végigvittnek, bár kétségtávol a hangzó és az írott nyelv különbségei diszciplináris módon leírhatóak. A diszciplináris vizsgálat rendszerint a lingvisztika (esetleg filológia) keretei között szokott zajlani, de ezen írás keretei között mégsem ezt a nézőpontot fogom érvényesíteni, hanem a leginkább etnolingvisztikainak nevezhető módszert, szemléletmódot. Az említett szemléletnek csak néhány mozzanatára lesz lehetőség kitérni, nincs szó teljességről vagy végig vitt rendszerességről. Számomra itt az etnolingvisztikai (vagy lingvisztikai etnográfiai) módszer elsősorban kulturális hagyományok, hétköznapi szentenciák tartalmának

elemzését és bizonyos szociolingvisztikainak mondható felvetéseket fog jelenteni.

### **Az írás: adomány, tehetség, kiváltság – kultúrhistóriai mozzanatok**

Az írásbeliséggel rendelkező beszédközösségek legtöbbje számot akar (és szokott) adni arról, hogy honnan is származik az írás, az írás tudása, tudománya. Minek, kinek köszönhető? Nem különbözik az efféle eredetfelvetés a hangzó nyelvre vonatkozó hasonló felvetésektől. A szanszkrit írásról – a szóban forgó kérdést megválaszolandó – használóik azt vélték, hogy írásuk az istenek városából való adomány. A régi egyiptomiak írásukról azt gondolták, hogy azt a bölcsesség istene ajándékozta nekik, és folytathatnánk a sort tovább. A lényeg – valószínűleg – az, hogy az írás egy olyan nagyszerű, ugyanakkor bonyolult dolog, hogy nem minősíthető egyszerűen emberi találmánynak.

### **Az írás hatalom**

Az ókori klasszikusok közül Plátón teszi az alcímben megfogalmazott kijelentést. Nagyon határozott megkülönböztetést tesznek tehát már az ókoriak a nyelv hangzó és írott formája között, a nagyobb presztízst az írásnak ítélve. Nehéz szabadulni másféle párhuzam felvetésétől is, de egy-két mozzanatra utalni kell, amely az érem másik oldalát mutatja, azaz az írástudatlanságot. Az írástudatlanság hátrány – több szempontból (erről majd később). Számunkra itt az a fontos, hogy az írás a szűkebben vett nyelvi információs folyamatokon túl meghatározott társadalmi funkcióval is fel van ruházva.

### **Hétköznapi – professzionális elmélet**

Az alcím többféle gondolat kifejtését igényelné, de itt csak kettő említésére vállalkozom. Mind a hétköznapi (laikus) ember, mind a szakember a világ jelenségeit értelmezni, magyarázni akarja, meg akarja érteni. Az egyes jelenségeket – ilyen vagy olyan módon csoportosítva – általánosítja, elméletet alkot, aminek lényege, hogy a világnak értelmet adjunk. E tekintetben a hétköznapi („népi”) és a tudományos gondolkodás között nincs radikális különbség, a törekvések hasonlóak, azonosak. Tehát a közvetlen tapasztalás, munkavégzés, társadalmi gyakorlat egyaránt nyer valamiféle absztrakciót akár tudományosat, akár „népit”. A személyes és a társadalmi élet számtalan területén tapasztalható ez a szembeállítás, de talán a legegyszerűbb a népi



gyógyászatra (természetgyógyászatra) és az orvosi gyógyászat különbségére hivatkozni.

Az orális és vizuális információszerzésre is vonatkozik a „szakmai” és a hétköznapi (laikus, „népi”) megközelítés és ennek nyomán az elméletalkotás. Nagyon sok nyelvi kifejezés utal arra a tényre, hogy az információszerzés és feldolgozás folyamata csak részben függ az információ fizikai megjelenítési módjától. Lépten-nyomon használunk olyan kifejezéseket, teszünk olyan megnyilatkozásokat, amelyekben a nyelvi kifejezés elsődleges (szó szerinti) jelentése konfliktusba kerül az aktuális jelentéssel. A jelenségnek többféle alosztálya van – mint majd az alábbi példák mutatni fogják (a rész utal az egészre, az aktuálisan használnak épp az ellenkezőjére kell gondolni stb.). Különböző nyelvekből idézhetnénk idevágó adatokat, de én csak néhány magyar példára szorítkozom, miközben ezeket is leszűkítem a *mond*, *olvas*, *beszél*, *ír*, *lát*, *ért* szavakra (velük kapcsolatban is csak néhány lehetséges példát idézek).

*Sokatmondóan nézett rám* – a megnyilatkozásban a *sokatmondóan* rész értelmezésében eltekintünk attól, hogy a *mond* ige elsődleges jelentésében az auditív-vokális csatorna jellegzetes tevékenységét igesíti, az aktuális jelentés azonban éppen ellenkezőleg: az információszerzés vizuális módját fogalmazza meg.

*Olvas vonásai között* – a vizuális információszerzés kulturálisan meghatározott módjának nyelvi kifejezésének (*olvas*) jelentéskörét tágitja, mintegy kiterjeszti azt a határt, ami a vizuális információszerzésben az *olvas* ige számára „természetes” módon ki van jelölve, tehát az alosztály tipikus jelentését osztály szintjére emeli.

*Világosan látja a helyzetet* – ez a megnyilatkozás a „rész utal az egészre” elvet érvényesíti. A *lát* ige elvileg csak a vizuális információszerzés és annak feldolgozási folyamatát jelenti, de aktuális jelentése a valóság nagyobb szeletére terjed ki. Általános észlelésre, értésre utal, függetlenül attól, hogy csak egy csatorna – a vizuális – szolgált az információ szerzésére, sőt maga a vizualitás a mondat jelentésének értelmezésekor nem is kap szerepet.

*Na látod!* – a megnyilatkozás szintén az előbbi elvet érvényesíti, az értés egészére vonatkozik függetlenül attól, hogy a *lát* ige a vizualitás jelentését hivatott kifejezni.

*Arcának nyugodtsága önmagáért beszélt.* – a megnyilatkozásban a *beszél* ige alapjelentése (oralitás) változik meg, és a vizuális információszerzés módjára utal.

*A szokás ilyenkor azt diktálja.* – megnyilatkozásban a *diktál* ige aktuális jelentése szintén nem az írott nyelvi tevékenység egyik szokásos alfajára utal. A megnyilatkozás jelentésének értelmezéskor eltekintünk a *diktál* ige alapjelentésétől.

*Kérem, írja le a helyzetet!* – a felszólítás teljesen rendjén van akkor is, ha például egy baleset történéseinek tisztázásaként hangzik el. A megnyilatkozást tevő rendőr egyáltalában nem gondol írásos tevékenységre. A *leír* ige aktuális jelentése hangzó nyelvi tevékenységre, azaz elmondásra, szóbeli beszámolóra utal. A közlő azt szeretné elérni, hogy a kérdezett a baleset körülményeiről (a szereplőkről, magukról az eseményekről, ezek pontos sorrendjéről stb.) adjon pontos tájékoztatást. (A példa azt is mutatja, hogy az elemzett jelenség, az orális és a vizuális úton nyert információ nyelvi kifejezésének át- és összefonódása nem tekinthető pusztán valamilyen nyelvtörténeti adaléknak, hanem kortól függetlenül a nyelvi magatartásunk jellemzője.)

A nyelvi példák számát természetesen szaporíthatnánk, de a lényeg így is kitűnik. A lingvisztikában olyan általánosan elfogadott, de valamennyi részletében még nem világos általános szemiotikai megközelítésről van szó, amely azt mondja ki, hogy a nyelvi üzenet (mint kód) lényegét nem befolyásolja a manifesztáció módja. Számunkra itt az a fontos, hogy az emberi beszédképesség – a maga legáltalánosabb törvényszerűségeit illetően – nem minősíthető eleve orális vagy írott formájúnak. Vonatkozik ez arra is, ha a nyelvi készségeket az innátizmus prizmáján át tekintjük. A modern nyelvi innátizmus egyik legismertebb alakja, N. Chomsky sem állítja egyetlen írásában sem, hogy a velünk született nyelvi készségek egyúttal hangzó nyelvi készségeket jelentenének. A fentiek alapján egyelőre azt kell leszögezni, hogy a hétköznapi nyelvszemiotikai szemlélet egy bizonyos szinten nem tesz éles különbséget az információt hordozó csatorna fizikai sajátosságai-ból adódó különbségek szerint. A jelet vevő és feldolgozó humán információs rendszert egynek és oszthatatlannak tekinti, az információt hordozó csatorna fizikai különbségeiből adódó információ-feldolgo-



zási sajátosságoktól kezdetben eltekint. Ez a „hétköznapi” (naiv) szemlélet egybeesik a „tudományos” közelítéssel.

### **Az írás szociolingvisztikája: nyelvökológiai szemlélet**

A nyelvökológiai szemléletből egy olyan fontos kijelentés is adódik, hogy az írás, az írott, a vizuális nyelv – mint tevékenység és mint eredmény egyaránt – a társadalmi környezet szerves része, a társadalmi tevékenység alkotóeleme. Csak néhány példa felvillantására fogok szorítkozni. A példák az írásbeliséggel rendelkező társadalmak, közösségek mindennapi életéből származnak, ez azonban nem jelenti azt, hogy a vizuális nyelv csak ezek körében játszik fontos szerepet.

### **Az írás mint korstílusjegy**

A művészettörténet és az írástörténet összefüggése jól ismert. A különböző művészi korszakok (reneszánsz, barokk stb.) nemcsak az építészetre, festészetre stb. vonatkozóan teremtik meg a maguk sajátos „formanyelvét”, hanem az írás elemeinek perceptuális és fogalmi szintjén is. Más-más betűtípus az uralkodó a gótika, a barokk stb. korszakban, de a szövegalkotás (az egyszerűség kedvéért nevezzük: fogalmi szintnek) tekintetében is jól kimutathatóak egy-egy korszak sajátos jegyei. Ebben az esetben az írott nyelv, mint a társalom, társadalmi tevékenység szerves része – mint említettem – perceptuális és fogalmi szinten számít környezetalkotó tényezőnek.

### **Írás és politika**

Az írás más vonatkozásban is szerves része a társadalmi gyakorlatnak. A görög demokrácia kiteljesedését többen is összekapcsolják az írásbeliséggel, annak fejlődésével és terjedésével. A dolog nem feltétlenül valamiféle mai ok-okozati összefüggésben tekintendő, azaz úgy, hogy a demokrácia az iskoláztatáson keresztül kedvez az írásbeliség terjedésének, mivel a lakosság (a polgárság) egyre nagyobb számban élvezi az iskoláztatás előnyeit, beleértve ebbe az írásbeliséget is. Inkább arról van szó, hogy a görög demokrácia egyes intézményesített eljárásai (például cserépszavazás) nem valósulhattak volna meg tömeges analfabétizmus feltételei és körülményei között. Így kell tehát értenünk az írott nyelv és a politikai fejlődés kapcsolatát: mindkét alkotóelem egyszerre játszhatja az ok és az okozat funkciót, valószínűleg bizonyos határokon belül kölcsönös az, hogy mindkettő feltételrendszert jelent a másik fejlődése, kiteljesedése szempontjából.



A fentebbi két példa jól mutatja, hogy az írott nyelv mikro- és makrotársadalmi környezeti feltétel és alkotóelem. Ezt a kettős funkcióját természetesen a modern társadalmakban is megőrizte.

A nyelv, mint mikrotársadalmi feltétel és szervező erő jelentkezik például az iskoláztatást megelőző gyermeki években is. Ha a gyermek közvetlen környezetét jelentő családban már az iskoláskor előtt az írott nyelv használatát tapasztalja (esti meseolvasás, szülők olvasási szokásai, kultúrája, levelezés a távolabbi rokonsággal stb.), akkor ez pozitív eredményeket prognosztizál az iskolai írás és olvasás tekintetében. Ha a gyermek számára az írás és az olvasás kizárólag az iskolába járással együtt kezdődő tevékenység, akkor legalább az újszerűség okozta kezdeti nehézségek valószínűsíthetők, de súlyosabb következmények is fenyegethetnek. Az írott nyelv, mint makrotársadalmi feltétel, nyelvpolitikai szerepe több szempontból is felvethető, én itt egy példára fogok szorítkozni. Ha kisebbségi nyelvterületen csak az állam nyelvén (ami példánk alapján eleve nem a kisebbség nyelve) függeszthetők ki különböző feliratok (utca-, bolt stb. nevek), akkor ez tulajdonképpen a kisebbségek nyelvi korlátozását jelenti, attól a jogtól való megfosztásukat, hogy nyelvük írott változata makronyelvi környezetének szerves része legyen.

Végül az írásbeliség, az írott nyelv ismeretének, elterjedtségének és a makrotársadalmi gyakorlat összefüggései közül említsük meg az analfabétizmushoz való viszonyt. A leginkább szokásos társadalmi magatartás ebben a tekintetben az orvosi szemléletet sugalló küzdelem az analfabétizmus ellen, amelyet mint egy járványt le kell győzni, fel kell számolni, illetve a társadalom ilyen irányú fertőzöttsége központosított beavatkozást igényel. Olykor ennél drasztikusabb megállapításokkal, majd ezekből fakadó szemléletmóddal is találkozunk. Azzal, amelyik az analfabétizmust a bűnözéssel, a társadalmi létezés alacsonyabb rendűségéből adódó egyéb hátrányokkal kapcsolja össze.

Végül az írásbeliség társadalmi, nyelvökológiai vonatkozásai közül azt említem meg, hogy az írásbeliséggel rendelkező társadalmakban többféle közlési csatorna alakul ki, és finomabban strukturálódnak az információközlési folyamatok. Gazdagabb lesz az információközlésért felelős személyi, tárgyi és helyzeti infrastrukturális háttér is. Az írásbeliség és annak birtokba vétele előnyösen alakítja át az egyéni kogni-

tív erőket, tágítja az egyéni kommunikációs szféra határait. Az írásbeliséggel rendelkező társadalmakban hátrányt jelent az analfabétizmus és annak „modern” változata a funkcionális analfabétizmus: a társadalmi munkamegosztásban eleve alacsonyabb rangú, kisebb presztízsű helyet képesek betölteni az írást egyáltalában vagy csak nagyon gyengén ismerők és használók.

### **Az írás pszicholingvisztikája**

Az előző alfejezetben az írás társadalmi összefüggései kerültek előtérbe. Ebben az alfejezetben a pszicholingvisztikai szempontnak megfelelően, egyéni, személyes jellegű mozzanatokat fogok említeni, természetesen nem a teljesség igényével.

### **Az írás kezdetei: ontogenetikus szempontok**

Az iskolába még nem járó gyermekek hol ilyen, hol olyan formában mutatnak érdeklődést az írás iránt. Ezek a próbálkozások rendszerint két nagyobb területet ölelhetnek fel, egyfelől az írott nyelvvel kapcsolatos metalingvisztikai ismereteket, másfelől a gyermeki rajzokat. A metalingvisztikai ismeretek – pontosabban: az ezek megszerzése iránti érdeklődés – legtipikusabb megnyilvánulásai a következők:

1. a fonéma-graféma megfeleltetési szabályok iránti érdeklődés (*Hogyan kell írni az Ákos betűt?*),
2. az első „betű” alapján a szavak kitalálása
3. a szavak szótagolása.

A fentebbi legtipikusabb gyermeki próbálkozások környezetfüggők: elsősorban azon gyermekek körében jelentkeznek, akik író-olvasó, az írott nyelvet különféle funkcióban használó családban nevelkednek. Könnyedén belátható, hogy ezek a gyermekek nem az iskolában ismerkednek az írott nyelv világával. Az iskoláztatás során a gyermekek közötti, gyakran lényegesnek bizonyuló különbségek az írott nyelv elsajátításában többek között e jelenségben is gyökereznek.

A gyermeki rajzok már kevésbé mikrotársadalmi (családi) függőségek. Rajzolni a legtöbb gyermek magától is rajzol: az óvodában, az aszfaltra, a homokba, a falra, tehát kevésbé van kiszolgáltatott helyzetben. A gyermeki rajzok több vonatkozásban jelenthetik az írott nyelv kezdeteit, de bennük korántsem minden megjelenő elemnek van jelentősége az írott nyelv szempontjából. A tartalmi tömörítés, egy-egy alak, forma köré kerített történet (jellemző vonás, mozzanat stb.) in-



kább az írott nyelv fogalmi, szövegszintű előfutárai. A gyermekrajzok grammatikája már az írást mint elvont jelrendszert vetíti fel. Ha tehát a gyermeki rajz viszonylag állandósult eszközök használatával képes kifejezni, hogy egy vagy több kutya van-e a rajzon, akkor már a gyermeki vizuális nyelv grammatikájáról beszélhetünk.

### **Az írás kiváltotta legfontosabb változások**

Az írás kevésbé előfeltételez, sokkal inkább meghatározott eredményekkel jár. Többnyire téves az az elgondolás, hogy az iskolai írástanulás például jó fonéma-megkülönböztető képességet előfeltételez, illetve, hogy csak ennek birtokában sikeres az írás megtanulása. Sokkal inkább arról van szó, hogy az írás során kényszerül rá a gyermek, hogy a szóban forgó műveletet (és az azt megalapozó, tehát lehetővé tevő részműveleteket) többnyire implicit módon megtanulja. A legtöbb gyermeknél ez a folyamat olyannyira gyors, hogy tulajdonképpen fel sem tűnik, ezért támadhat az a benyomás, hogy a fonéma-megkülönböztető képesség feltétel, nem pedig eredmény. Ha ez a rejtett tanulási folyamat túlságosan elhúzódik, és többszöri tisztázást igényel, azaz nem zökkenőmentes, akkor rendszerint ez alapot ad a cserére: feltételként jelenik meg az, ami funkcióját tekintve eredmény lenne.

A fonéma-graféma (graféma-fonéma) megfeleltetési szabályok megtanulása olyan ismeretekhez juttatja a gyermeket, amelyekre az írást nem ismerő gyermek nem vagy csak erősen közvetve tesz szert. A szóban forgó szabályok megtanulása során a gyermek kénytelen pontosan kielemezni a hangalaki megvalósulások mögötti állandóságot, felfedezni a variánsok között az invariánst, azaz azt, hogy például a *meleg+hez*, *meleg+ben* szavakban a grammatikai morféma előtti szegmens a változás ellenére is ugyanaz. Ez az ismeret morfémakonsztanciát, a morfémahatárok felfedezését biztosítja és eredményezi. Implicit módon tudatosul a szavak morfematikus felépítése, szerkezete. Összefoglalóan azt mondhatjuk, hogy alaposan megváltozik az írott nyelvet is birtokló idiolektus. E tekintetben az írni tanuló gyermek tevékenysége, feladata és teljesítménye nem különbözik a nyelvelsajátítás korábbi szakaszától, azaz az első életévek orális nyelvi elsajátításától. Az írásra is kiterjeszthetjük: a gyermek kis nyelvészként kell, hogy viselkedjen és járjon el.

### **Az írás: filogenetikusszemponatok**



Az írás pszicholingvisztikai irányultságú vizsgálatát nem segítette az a részigazság, hogy az írás filogenetikusan a memória támogatását szolgálta. A Korán és más szent írások recitálása, terjedelmes népköltészeti alkotások, eposzok szájhagyomány útján való fennmaradása mind-mind az említett kijelentés félígazságát hangsúlyozza (azt itt nem részletezzük, hogy a szóban forgó alkotások verbatim, formahű megőrzését sajátos nyelvi eszközök – ritmika, tartalom stb. – hathatósan segíti).

Azt, hogy az írás filogenetikusan nemcsak a memória támogatására szolgált még hathatósabban támogatja az ókori feliratok egy része. A mezopotámiai, egyiptomi, hettita, görög és római feliratok egy része néhány szavas, soros üzenet az olvasónak egy adott uralkodó dicsőségéről. Ezek a feliratok (az írás ellenére!) tulajdonképpen nem olvasót, hanem hallgatót feltételeznek, azaz a címzett nem az olvasó, hanem a hallgató ember. Bizonyos szövegekkel egyébként ma sincs ez másképpen.

Az ókori királyokat dicsőítő feliratok alkotói aligha az emberi memóriakapacitás szűkössége miatt választották ezt a módszert. Sokkal inkább arról a – ma is elfogadott és gyakorolt – álláspontról van szó, hogy az írás az emberi nyelvi kommunikáció „erősebb” formája. A két forma – hangzó és írott – tehát ugyanazt a lényeget fejezi ki, de meghatározott funkcióban az egyik a másiknál erősebb, hatékonyabb. E kettősség, mint dióhéjban felvillantva látszik, már a történelem hajnalától fogva él.

### **A témához kapcsolódó irodalom**

Baranyi Erzsébet–Lénárt Edit: Az írásbeli közlés gondolkodáslélektani vonásai. Akadémiai Kiadó. Budapest 1959.

Barton, D.: An Introduction to the Ecology of Written Language. Blackwell. Oxford 1994.

Besner, D.–Humphreys, G. W. (eds.): Basic Processes in Reading. Visual Word Recognition. Lawrence Erlbaum Ass. New Jersey 1991.

Black, M.–Byng, S.: Prosodic Constraints on Access in Reading. Cognitive Neuropsychology 3. 1986, 369-409.

Doblhofer, E.: Jelek és csodák (Letűnt írások és nyelvek megfejtése). Gondolat. Budapest 1962.

Elekfi László: Fonetikus írás. In: Nyelvművelő kézikönyv I. Szerk.: Grétsy László–Kovalovszky Miklós. Akadémiai Kiadó. Budapest 1980, 664-668.

B. Fejes Katalin: Egy korosztály írásbeli nyelvhasználatának alakulása. Tankönyvkiadó. Budapest 1981.

Fekete Péter–Raisz Rózsa (szerk.): A szöveg szerkesztése, megértése, kidolgozása és megszólaltatása. Országos Anyanyelv-Oktatási Napok. Eger 1993.

Gorrell, R.–Brown, M. M.: Writing and Language. Harcourt Brace Jovanovich. New York 1972.

Goswami, U.–Bryant, P.: Phonological Skills and Learning to Read. Lawrence Erlbaum Ass. New Jersey 1990.

Gósy Mária: Szókeresés a mentális lexikonban. Magyar Nyelvőr 122/2. 1998, 189-201.

Halliday, M. A. K.: Spoken and Written Language. Oxford University Press. Oxford 1985.

Lengyel Zsolt: A fonetikai és a fonológiai elv az írott nyelv ontogenezisében, In.: Beszédkutatás '97 (Szófonetikai vizsgálatok) Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1998, 118-142.

Lengyel Zsolt: Az írás: kezdet – folyamat – végpont. (Az írástanulás pszicholingvisztikai alapjai). Corvina. Budapest 1999.

Ligeti Róbert: Az írástanulás pszichológiája. Budapest 1982.

Nystrand, M.: The Structure of Written Communication (Studies in Reciprocity between Writers and Readers). Harcourt Brace Jovanovich. New York 1986.

Papp Ferenc: Írástanítás – a nyelvész szemével, In: Fülöp K.–Lengyel Zs. (szerk.): Az orosz nyelv tanítása az általános iskola 4. osztályában. Szeged 1982, 147-165.

Reeder, K.: Text or Context? The Influence of Early Literate Experience Upon Preschool Children's Speech Act Comprehension. In: Conti-Ramsden, G.–Snow, C. E. (eds.): Children's Language. Vol. 7. Lawrence Erlbaum Ass. Hillsdale, New Jersey, Hove and London 1990, 305-327.

Scarcella, R.–Lee, Ch. Different Paths to Writing Proficiency in a Second Language? A Preliminary Investigation of ESL Writers of Short-term and Long-term Residence in the United States. In Eisenstein, M. R. (ed.): The Dynamic Interlanguage (Empirical Studies in Second Language Variation). Plenum Press. New York 1989, 137-155.

Scinto, L. F. M.: Written Language and Psychological Development. Harcourt Brace Jovanovich. New York 1986.

# **A MENTÁLIS LEXIKON REPREZENTÁCIÓJA ÉP ÉS BROCA-AFÁZIÁS SZEMÉLYEKNÉL**

**Hoffmann Ildikó**  
**MTA Nyelvtudományi Intézete**

## **Az afázia**

Ha sérülés éri az agy valamely részét, annak bizonyos funkciói megváltozhatnak. A jelenségnek beszédbeli visszatükröződése is lehetséges. A nyelv és a beszéd központi idegrendszeri meghatározottságának vizsgálatánál az afázia került a központba. Az afázia: zavar a beszédpercepcióban (az észlelésben és a megértésben) és/vagy a beszédprodukcióban, s az agy szerzett organikus sérülésével függ össze. A leépülés mindig a már elért beszédszinthez képest történik.

Nyelvi szempontból nem beszélhetünk két teljesen azonos típusú afáziás betegről. Előfordulhat, hogy két beteg afáziájának okozója ugyanannak az agyi területnek a sérülése lesz, nyelvhasználati zavarai mégis különbözőek.

Az afázia típusai közül tanulmányunkban Broca- (agrammatikus) afáziás betegek és különböző életkorú felnőttek mentális lexikonhoz való hozzáférését vizsgáljuk. A Broca-afázia során a motoros területet éri a sérülés. A beszédpercepció jól megtartott, bár megértésbeli problémák a mondat szintjén felmerülhetnek. Súlyosan érintett a beszéd expresszív része, azaz a beszédprodukció. Morfológiai szinten toldalékvesztés, szintaktikai szinten pedig a produkált mondatok hosszúságának csökkenése, a szókészlet korlátozódása, az agrammatizmus jelensége tapasztalható. A sérülés mind a funkciószavakat, mind a tartalmas szavakat érinti (Bánréti 1995). E két szóosztály előhívásának deszinkronizációja az agrammatizmus egyik forrása. Az agrammatizmus okait vizsgálva további két feltételezés merült fel: a) az alapvető probléma a beszélt nyelv szerkezeti zavara, b) a szerkezet ép, a probléma a feldolgozásbeli korlátozottságban keresendő, amelynek felelőse a munkamemória.



Az agrammatizmus vizsgálatánál különböző tesztelési módszerek léteznek. Ilyen, részképességeket mérő, a magyar nyelvre is kidolgozott parciális tesztek: a) a grammatikalitást megítélő, b) a mondatismétléses (Bánréti 1995, 1997), valamint c) a mentális lexikonhoz való hozzáférést vizsgáló tesztek.

### **A mentális lexikonhoz való hozzáférés vizsgálata**

#### **A mentális lexikon szerveződése**

Csaknem valamennyi afáziás beteg küzd bizonyos mértékben megnevezési zavarral<sup>1</sup>, amely a belső lexikonhoz való hozzáférés zavara (Osmanné Sági 1995). Feltételezhető, hogy a sérülés során a szavak nem törölődnek ki a memóriából, hiszen az adott szót bizonyos szituációban megtalálja a beteg, más esetben azonban előfordulhat, hogy nem. A lexikon hozzáférési útja lesz tehát az, ami zavart szenved.

A mentális lexikon az összes olyan információt tartalmazza, amely egy adott nyelv megszerkesztéséhez szükséges. Szerveződésére számos hipotézist találunk. Az afázia irodalmában egyre többet olvashatunk egy olyan módszerről, amely a lexikai aktiváció időbeli paramétereit kutatja: a lexikai előfeszítési, támogatási hatásról (priming-hatás). Ennek a módszernek a tesztelési eredményei megerősítik azt az állítást, amely szerint a lexikai aktivizálásnál a szemantika az, amely a leghamarabb működésbe lép. A lexikális hozzáférés folyamata a beszédprodukciónál a következő: a szükséges lexikai egységet kiválasztjuk, fonológiai/fonetikai formába öntjük, majd az aktuális artikulációs gesztusok véghezvitelével kimondjuk. A beszédmegértésnél akusztikai jelsorozatot fogadunk, hozzárendeljük a megfelelő fonetikai jegyeket, azonosítjuk a lexikonban tároltakkal, s – ideális esetben – megtaláljuk a keresett egységet (Levelt 1989; Gósy 1998).

---

<sup>1</sup> A megnevezési zavar károsodása: anómia. Osmanné Sági (1991) négy fő típusát különbözteti meg: szóprodukción anómia (ezen belül artikulátoros iniciáció zavara és parafáziás anómia), szóselekciós anómia (tisztá szótalálási zavar), szemantikus anómia és diszkonnekciós anómia (modalitás-specifikus és kategória-specifikus).

A mentális lexikont sajátos és gyors időstruktúra jellemzi. A folyamatos beszéd létrehozása és megértése során döntően komplex, párhuzamos folyamatok mennek végbe. Valószínűleg kevesebb időre van szükség a szemantikailag szorosabban kapcsolódó egységek aktivizálásához, mint a nem túl szorosakéhoz. A mások által irányított aktivizálás – így a teszhelyzet – nehezítő körülmény lehet.

### **Hipotézis, korpusz és tesztelési módszerek**

Kiindulól hipotézisünk az, hogy annak függvényében, milyen úton próbálunk előhívni szavakat, más-más eredményességet kapunk, azaz a különböző tesztelési módszerek ugyanazon vizsgálati személynél a mentális lexikon reprezentációjának eltérő eredményességével zárulnak. Ahhoz, hogy adott személy, illetve afáziás beteg belső lexikonjáról globális képet kaphassunk, szükséges több lehívási utat is bejárni. A vizsgálatunkban három csoport vett részt:

- **Broca-afáziás csoport** (6 fő), átlagéletkor 60 év;
- **I. kontroll csoport** (20 fő), átlagéletkor 25 év;
- **II. kontroll csoport** (20 fő), átlagéletkor 60 év.

Az összehasonlító eredmények hitelessége tekintetében fontosnak tartottuk egy fiatalabb korosztályból álló kontrollcsoport és egy, a Broca-afáziás csoporttal azonos átlagéletkorú normál/egészséges csoport vizsgálatát.

A mentális lexikonhoz való hozzáférést kutatva olyan teszteket vetünk figyelembe, amelyek különböző nyelvi szintek bevonásával vizsgálják a jelenséget. A következő négy tesztet választottuk ki:

1. képmegnevezés (tárgy és cselekvés),
2. mondatkiegészítés (főnév és ige),
3. zárt eljárás (főnév, ige, névmás, határozószó),
4. szófolyamatossági teszt (főnév és ige).

### **Kísérleti eredmények**

A képmegnevezés, a mondatkiegészítés és a zárt eljárás tekintetében a két normál kontrollcsoport 96-100%-os teljesítményt nyújtott. Eltérően alakultak az eredmények a szófolyamatossági tesztelés során. Ennek

adatait a Broca-afáziás csoporttal együtt ismertetem. Következzenek tehát az afáziás betegek tesztelési eredményei.

### 1. Képmegnevezés

A megnevezés általános modellje a következő: vizuális analízis → szemantikai aktiváció (kategória kijelölés, lexikai egység kitöltése) → a keresett szó fonológiájának aktivációja, reprezentációja /artikulált beszéd/ (Levelt 1989; Gósy 1998). A vizsgálatunkban felhasznált teszt a Boston Naming Test (Goodglass–Kaplan–Weintraub 1983) és egy cselekvéses teszt volt. A Boston Naming Test 60 darab fekete-fehér kontúrrajzot és egy értékelőlapot tartalmaz. Az általunk felhasznált másik tesztben 30 cselekvés képét mutattuk meg. A vizsgálati személyek feladata a kísérlet vezetője által mutatott képen lévő tárgy vagy cselekvés megnevezése volt.

A hat Broca-afáziás vizsgálati személy átlagteltjesítménye a feladatmegoldások tekintetében 72%-os volt. Ahol problémák voltak a szótalálással, a betegek kompenzációs stratégiaként a főneveknél körülírtak, egyed- helyett kategórianévet hívtak le (például *ananász* helyett *gyümölcs*), az igénél pedig vagy csak a bővítményt, vagy bővítménnyel együtt neveztek meg. Számos esetben azonban a keresés sikertelen volt, a szótalálás meghiúsult.

A képmegnevezéses teszt maga nem tisztán nyelvészeti – a képi-vizuális analízis a jobb félteke bevonásával a bal félteke szótalálási műveletében segítő lehet. A tesztelés során nyújtott jó eredmények nem adnak tehát okot arra, hogy a vizsgált afáziás csoportban enyhe zavart feltételezzünk.

### 2. Mondatkiegészítés

A mondatkiegészítés a szintaxis szintjére visz el. Két feltevés adott arra vonatkozóan, hogyan használjuk fel a lexikai elemek által szerzett információkat (Friederici 1995). Az első hipotézis szerint struktúra-vezéreltségről beszélünk, azaz az elemző az első lépésben azonosítja a bejövő (hallott vagy olvasott) információk szerkezetét a szintaktikai szókatégoriák alapján, majd megtörténik a tematikus szerepek kiosztá-



sa, s felhasználásra kerülnek a lexika és a diskurzus információi. A második feltételezés szerint lexikai vezéreltségről van szó. Az elemzött a szintaktikai szerkezet helyett a lexikai információk irányítják, amelyek elérhetők már a szerkezet kiépítése előtt. A mondatkiegészítés során valamely mondatbeli lexikai információnk hiányzik. A mondat szerkezet segítőt vagy éppen nehezítőt lehet a keresett szó megtalálásában. Problémát okozhat az afáziás betegeknek a félreolvasás és a mondatok lineáris szerkezetépítésére, megértésére való törekvés, mely azután agrammatizmushoz vezet.

A vizsgálati feladatunk 10-10 mondatból állt. A mondatokban adott helyen egyetlen szó volt kihagyva. Az üres helyek részben ígét, részben főnevet jelöltek. A mondatok között szerkezetüket tekintve voltak egyszerűek és összetettek. A kihagyott szónak a szófajiságán és a mondatbeli szerepén kívül a lehvívás szempontjából az is fontos volt, hogy a mondat mely részén helyezkedett el. A mondat első szavának behelyettesítése mindig nehezebb feladatnak bizonyult.

A feladat elvégzése során megkértük a beteget, hogy hangosan olvassa fel a kiegészítendő mondatot. Időbeli korlátozást nem használtunk. A Broca-afáziás csoportnál a főnevek megtalálása gyorsabban, könnyebben ment, kevesebb volt a tévesztés. Gyakori volt a névelő elhagyása nemcsak mondatkezdetnél, hanem a mondaton belül is. A főnevek eseteit általában helyesen választották ki, de előfordult az alany eset felcserélése tárggyal (például *térkép* helyett *térképet*). Ha az adott szó nem jutott a beteg eszébe, körülírást alkalmazott (*térkép* helyett *részletesen megrajzolt kép*). Néhány esetben az üres helyre kiválasztott kategóriánév jó volt, a specifikus megnevezés azonban nem, de a jelenség fordítottjára is találtunk példát (*gyógyszert* helyett *kavintont*, *az oroszlán* helyett *a leopárd*, *a szobrász* helyett *Gábor*). A megnevezés a megnevezendő valamely tulajdonsága alapján történt a következő példákban: *telefon* helyett *csengő*, *berregő*; *béka* helyett *brekecs* (hangutánzás). A téves/rossz megoldások agrammatikus mondatot eredményeztek (*Hiába csengett hosszasan a telefon, senki sem vette fel.* helyett *Hiába csengett hosszasan Petike, senki sem vette fel.*). Az ígék megtalálása lényegesen hosszabb időt vett igénybe. Előfordult, hogy az igekö-

tős ige helyett csak az igekötőt helyettesítette be a beteg az üres helyre (*leesett* helyett *le*). Megfigyelhető jelenség volt a jó argumentumszerkezet megtalálása mellett a szemantikailag helytelen ige lehívása. További megfigyeléseink: az ige és az igekötő felcserélése (*ígérte meg* helyett *megígérte*), bővítménnyel együtt való megnevezés (*kergette* helyett *kergette az udvaron*), nem releváns ige kiválasztása (*sütött* helyett *lekattintotta*), téves ragkiosztás (*szólt* helyett *szóltak*) és az ige helyett más szófaj lehívása, mely ugyanakkor nem befolyásolta a mondat grammatikailagosságát (*A gyümölcs megéri a naptól.* helyett *A gyümölcs rózsaszínű a naptól.*).

Mindezeket összesítve megállapíthatjuk, hogy a két szófaj, a főnév és az ige között a tévesztések és a helyes szótalálások tekintetében nem volt kimutatható szignifikáns különbség. Fonémikus parafáziák mindkét szófaji kategóriában közel azonos arányban fordultak elő (*virág* helyett *viseg*, *viszág*; *hullámozott* helyett *hullámozot*; *érkeztek* helyett *énetek*, *énekek*).

### 3. Zárt eljárás

Az eljárás a kontextus kényszerítő erejét, s így a szótalálásban való szerepét is vizsgálja (Lengyel 1995). A vizsgálati anyagunk: egy rövid történet (kétszereplős mese), előre meghatározott helyeken szókihiagyással. A kitöltendő szavak száma tizenkettő. A vizsgálati személynek a történet egyszeri átolvasása után az újraolvasáskor az egy-egy hiányzó helyre azt az egyetlen szót kell beírnia, amelyet a szövegkörnyezetben a legodailobbnek ítél.

A nyelvi viselkedés kontextusfüggő. A szövegösszefüggéssel kapcsolatos feladatok teljesítésének feltétele összetett: a szintaktiko-grammatikai folyamatok mellett szemantiko-pragmatikaiak is lejátsszónak (Lengyel 1995).

Az afáziás betegek üres hely kitöltésének típusai a pontos behelyettesítés mellett a következők voltak: a) azonos szófajú, grammatikailag és szemantikailag helyes szócsere, b) azonos szófajú, grammatikailag helyes, szemantikailag helytelen szócsere, c) azonos szófajú, grammatikailag helytelen, szemantikailag helyes szócsere, d) azonos szó-



fajú, grammatikailag és szemantikailag helytelen szócsere, e) más szófajú, grammatikailag és szemantikailag helyes szócsere, f) más szófajú, grammatikailag és szemantikailag helytelen szócsere és g) nincs válasz.

A teszt kiértékelése során megkülönböztettünk „pozitív kategóriát” (pontos behelyettesítés; azonos vagy más szófajú, grammatikailag és szemantikailag helyes szócsere) és „negatív kategóriát” (azonos szófaji kategória, grammatikailag helyes, szemantikailag helytelen szócsere; azonos szófajú, grammatikailag helytelen, szemantikailag helyes csere; más szófajú, grammatikailag és szemantikailag is rossz csere; kitöltetlen hely). A hat Broca-afáziás beteg egymástól is eltérő eredményességgel oldotta meg a tesztet. Az üres helyek pozitív és negatív kitöltési megoldásait szemlélteti a következő összefoglaló táblázat (Hoffmann 1999).

1. táblázat: Megoldások a Broca-afáziás csoporton belül

Kategória/Személy	M. I.	M.G.	D. J.	K. Z.	K. S.	F. M.
Pozitív (%)	91,6	83,4	82,3	100	50,1	16,6
Negatív (%)	8,4	16,6	17,7	-	49,9	83,4

Az ige–főnév szótalálási oppozíció tekintetében a következő megállapításokat tehetjük. A kitöltési helyek főnév–ige aránya 6:3-hoz volt. A két fő kategórián kívül két névmás és egy határozószó üres helyét is be kellett helyettesíteni. A legtöbb helyes kitöltés a főnévnél fordult elő, amelyet azután az ige, majd a névmás követett. Legnehezebbnek a határozószó megtalálása bizonyult.

Az afáziás betegek szótalálási problémái a kontextus szintjén a korlátozott munkamemória kapacitásukból is adódnak. Megfigyelhető jelenség, csakúgy, mint a mondatkiegészítésnél, hogy a betegek itt is a linearításra törekednek egy-egy mondat megértésénél. A félreolvasások ebben az esetben is téves megoldásokhoz vezetnek.



#### 4. Szófolyamatossági teszt

A szófolyamatossági tesztben szintén valamely előre meghatározott szempont szerint kell aktivizálni a mentális lexikont. A teszt három részből áll. A vizsgálati személy feladata az, hogy adott időegységen belül a vizsgálat vezetője által megadott kezdő hanggal először főneveket, másodszorra egy másik kezdő hanggal igéket, harmadszorra pedig egy újabb hanggal szófaji megkötöttség nélkül soroljon fel szavakat. A vizsgálatunkban felhasznált hangok: [ p, k, b ]. A feladat megoldásához rendelkezésre álló idő minden egyes alkalommal másfél perc volt. A feladatok között a két kontrollcsoportnál 5 perc szünetet tartottunk, az afáziás csoportnál ennél többet, 15-20 percet.

A fennálló elmélet szerint (Gósy 1998) a legkönnyebb annak a feladatnak a teljesítése lesz, ahol nem határozunk meg szófaji megkötöttséget. A megtalált elemszámok tekintetében ezután a főnevek következnek, s végül az igék. A sikeresen előhívott főnevek és igék számában szignifikáns különbség nem tapasztalható.

Mennyiben igazolható ez az állítás az agysérülést szenvedett betegek körében, jelen esetben a Broca-afáziás betegeknél? A következő összefoglaló táblázat azt mutatja be, hogyan alakultak a szófolyamatossági teszt eredményei a vizsgálatunkban részt vett három csoportnál.

2. táblázat: Elemszám-adatok a szófolyamatossági teszt alapján

Vizsgálati csoportok	Főnév (db)	Ige (db)	Szó (db)
Broca-afáziás csoport (átl. életkor: 60 év)	6,3	5	8,3
I. kontroll csoport (átl. életkor: 30 év)	23,33	23	27,33
II. kontroll csoport (átl. életkor: 60 év)	17	16,66	23,66

A 2. táblázatból kiderül, hogy az elmélet igazolást nyert a Broca-afáziás csoportban is. A betegek a legtöbb szót a szófaji megkötöttség nélküli feladatban aktivizálták, majd a főnevek, s végül az igék következ-

tek. A főnevek és az igék megtalálása között itt sem találunk szignifikáns különbséget.

A táblázatot tovább tanulmányozva látható, hogy érdekesen alakultak az eredmények a szófolyamatossági tesztben az elemszámok tekintetében is. A Broca-afáziás betegek lényegesen gyengébben teljesítettek, mint a másik két egészséges csoport, s ez már erős szótalálási zavart jelez.

Ha szemügyre vesszük a két kontrollcsoport szóaktivizálási rátáit, itt is különbségeket tapasztalhatunk. A II., azaz a 60 éves átlagéletkorú csoport szignifikánsan gyengébben teljesített, mint az I., azaz a 30 éves átlagéletkorú vizsgálati személyek mind a főnevek, mind az igék és a szavak megtalálása tekintetében is.

3. táblázat: A szóaktivizálási teljesítmények százalékarányai

Vizsgálati csoportok	Főnév (%)	Ige (%)	Szó (%)
I. kontroll csoport (átl. életkor: 30 év)	100	100	100
II. kontroll csoport (átl. életkor: 60 év)	72,8	72,4	86,5
Broca-afáziás csoport (átl. életkor: 60 év)	27	21,7	30,7

A kapott eredmények tehát azt mutatják, hogy a mentális lexikonhoz való hozzáférés nemcsak agysérülés-, hanem életkorfüggő is.

Az ige–főnév oppozíció mindkét kontroll csoportban igazolódott, vagyis az, hogy a szótalálásban nincsen közöttük szignifikáns különbség. A szó kategóriában a főnevek domináltak, majd az igék, s végül a melléknevek következtek.

A szavak aktivizálása során valamennyi vizsgált csoportbeli egyén szakaszos szüneteket tartott. A Broca-afáziás betegek lényegesen hosszabb szüneteket produkáltak, mint a két kontrollcsoportbeliek. Az egy-egy szó közötti hosszabb szünet után az előhívásnak általában valamiféle stratégiáját próbálták a vizsgálati személyek követni. A két normál

csoportban leggyakrabban a kezdő szótag (például *baka, bab Baranya; károg, kántál; pap, pala, pad, park; palack, palánk, paré*) és a szemantikai összefüggés (például *piszok, portölő, porszívó; bárány, béget; párna, paplan; keres, kutat; poszter, plakát*) alapján hívódott le a következő elem. A fiatalabb kontrollcsoportban szerepe volt még a hangzásbeli hasonlóságnak is (*banda, bandukol; pinty, pint, pont*). Az idősebb kontrollcsoport tagjainál sokkal gyakrabban bukkant fel a véletlenszerűség (*kíván, küldet, kapkod; kapál, kúszik, kér; béke, bendő, barna*) és a szótagszámtól való függés (*por, per, pír; púp, pép; banya, bányá*). Az afáziásoknál a kezdő szótag mellett (például *papír, Panni, Pali; pipa, pilóta*) a véletlenszerűség dominált (például *kóдол, kapál; padló, porcelán, puttony; baba, beszéd*).

### Következtetések

A vizsgálatunk eredményei bizonyították, hogy a mentális lexikonhoz való hozzáférés különböző útjai más-más eredményességgel működnek. Broca- (agrammatikus) afáziában a szavak lehívása zavart. A lexikonhoz való hozzáférés tesztfüggő, amely azt jelent, hogy bizonyos tesztek alapján enyhe zavar feltételezhető (képmegnevezés), más esetek azonban súlyos zavart reprezentálnak (szófolyamatossági teszt). Kiinduló hipotézisünk tehát igazolódott.

Az afáziás csoporton belül egyéni eltéréseket is regisztráltunk. Az ige-főnév szófaji oppozíció tekintetében elmondható, hogy igékhez nehezebb a hozzáférés, a keresési folyamat hosszabb időt vesz igénybe. A főnevek megtalálásához valamennyivel rövidebb idő szükséges, s a sikertelen keresés esetén a beteg kompenzációs stratégiaként a körülírás-hoz fordul. A feladatokhoz felhasznált idő tekintetében a keresési műveletek a „normál” idő többszörösei voltak. Ez a tény már önmagában is jelzi a zavart.

A két, eltérő életkorú kontroll csoport tesztelésének eredményei a képmegnevezés, a mondatkiegészítés és a zárt eljárás során megegyeztek. Szignifikáns különbséget tapasztaltunk azonban a szófolyamatossági tesztelés alkalmával. Tudjuk, hogy a mentális lexikon döntően az anyanyelv-elsajátítás során alakul ki, de életünk végéig változik, módo-



sul (Gósy 1998). Vizsgálatunk további bizonyítékul szolgálhat arra, hogy a lexikon változása, az elemekhez való hozzáférés életkorfüggő, azaz az idős korra a szavak lehívása nehezebbé válik, hosszabb időt vesz igénybe, kevesebb elemszám produkálható. A munkamemória kapacitásának korlátozódásával itt is számolnunk kell.

A tanulmányban ismertetett vizsgálat teszt sorozat volt. A teszt szituáció blokkoló hatású lehet. A mentális lexikon reprezentációjának kutatásakor éppen ezért fontos a természetes körülmény, a spontán beszéd vizsgálata is, hiszen a szótalálási/visszakeresési zavarok itt is jól nyomon követhetők.

Befejezőként utalni szeretnénk a nyelvészeti afáziakutatás fontosságára, jelentőségére. A hazai afáziakutatás még fiatal tudományág. Valamennyi empirikus kutatáson és összehasonlító vizsgálatokon alapuló eredmény adalékként szolgálhat ahhoz, hogy a már meglévő nyelvi modellek helytállóságát igazolhassuk, esetlegesen cáfolhassuk, s pontosabb ismereteket szerezhessünk olyan nyelvi folyamatokról, mint például a mentális lexikon működése. Az afázia tudományától ugyanakkor nem választható el a terápia fogalma: valamennyi eredmény segítséget nyújthat a sikeres rehabilitációban, abban, hogy egy egyik pillanatról a másikra elveszett képesség, a beszéd és ennek megértése újra birtokunkba kerülhessen.

### Irodalom

Bánréti Zoltán: Modulok a mondatfeldolgozásban. In: A mai magyar nyelv leírásának újabb módszerei. Szerk.: Maleczki Márta–Büky Béla. Néprajz és Nyelvtudomány XXXVI., 1995, 109–128.

Bánréti Zoltán: Megjegyzések a neurolingvisztikáról. In: Bánréti Zoltán (szerk.): Nyelvi struktúrák és az agy. Neurolingvisztikai tanulmányok. Corvina. Budapest 1999.

Friederici, A.: The Time Course of Syntactic Activation during Language Processing: A Model Based on Neuropsychological and Neurophysiological Data. *Brain and Language* 50. 1995, 259–281.

Goodglass, H.–Kaplan, E.–Weintraub, S.: The Boston Naming Test. Lea–Fibiger. Philadelphia 1983.

Gósy Mária: Szókeresés a mentális lexikonban. *Magyar Nyelvőr* 122/2. 1998, 189–201.

Hoffmann Ildikó: Kényszerít-e a kontextus? *Medicina et Linguistica* IV. Budapest 1999, 103-109.

Lengyel Zsolt: Alkalmazott pszicholingvisztika. Egyetemi Kiadó. Veszprém 1995.

Levelt, W. J. M.: *Speaking*. MIT Press. Cambridge 1989.

Osmanné Sági Judit: Az afázia klasszifikációja és diagnosztikája I. Az osztályozás elméleti alapjai és a Western Aphasia Battery felépítése. *Ideggyógyászati szemle* 44. 1991, 339-350.

Osmanné Sági Judit: A belső lexikonhoz való hozzáférés zavarai afáziás betegeknél. *ÁNyT* XVIII. Akadémiai Kiadó. Budapest 1995, 147-163.

# A SZÓKEZDET SZÓKIVÁLTÓ HATÁSÁNAK VIZSGÁLATA AFÁZIÁS BETEGEKNÉL

Osmanné Sági Judit  
MTA Pszichológiai Intézete

## Bevezetés

A szóprodukciónban a mentális lexikon két szintjének (a lemma és a lexéma szint) megkülönböztetése jelenleg többnyire elfogadott (vö. Gósy 1998; Levelt et al. 1999). A lexémák a szavak fonológiai mintázatát tartalmazzák, míg a lemmák a szemantikus vonásokat, illetve szintaktikai információt (bár ez utóbbi erősen vitatott, vö. Caramazza 1997).

Az afáziás betegek többségénél a szótalálási zavar általános tünet. Változó lehet azonban, hogy a beszédképzés mely szintje károsodik. A két szint megkülönböztetésében igen gyakran segítséget adnak azon módszerek, amelyek a szóforma valamely fonológiai jellemzőjével próbálják aktiválni a fonológiai kimeneti lexikont (például ritmus, szótagszám). Nyilvánvaló, hogy ha ezek az eljárások segítséget adnak a szavak előhívásában, akkor azok lemma szintje ép, sőt a betegek a szóformáról is rendelkeznek némi információval. Az is lehetséges viszont, hogy a lexéma szint is ép, de a posztlexikális tervező szakaszok károsodtak (Goodglass et al. 1976). A fonológiai segítő információ-típusok közül a leginkább alkalmazott és talán a leghatékonyabb a szókezdés, amikor a vizsgáló személy vagy terapeuta megnevezi az első szótagot, egy szótagos szavaknál a kezdő hangot. Nem kívánjuk itt összefoglalni a szókezdet kiemelt szerepét a mentális lexikonban (vö. Marslen-Wilson cohort-elmélete, 1989). A szókezdés (a továbbiakban fonémikus támpontnak nevezzük) sajátosságait Goodglass és munkatársait követően szisztematikusan alig vizsgálták, ahogy ezt Caplan igen alapos neurolingvisztikai munkái is tanúsítják (1992).

A klinikai vizsgálatokban jól ismert, hogy a Broca-afáziás betegeknél (akiknél elsősorban a szintaktikailag teljes mondatok képzése



károsodik, de szótalálási nehézségekkel is küszködnek) a fonémikus támpont sikeresebb a szavak aktiválásában, mint a fluens afáziás betegeknek (ez utóbbiaknál ép a beszéd szintaktikai szerveződése, viszont dominálnak a beszédmegértési, illetve a morfológiai zavarok). Vizsgálni kívántuk, hogy e sajátosság kimutatható-e kontrollált körülmények között is – ilyen az egyszerű képek megnevezése.

A feladat maga azt feltételezi, hogy a tárgy képe legalább részben aktiválja a tárgy lexikális mentális reprezentációját, szemantikus mezőjét, azaz a lemma szintet.

A megnevezési zavar több forrásból adódhat. Ezek közül a leggyakoribbak az általános vizuális észlelési zavar, a vizuális agnózia, a vizuális reprezentáció és a verbális lemma szint közötti diszkonnekció, a lemma szint és a lexéma szint közötti zavar, a lexéma szint aktivációjának vagy magának a lexéma szintnek a zavara, illetve a posztlexikális beszédtervezési szakaszok zavara.

A szókezdet szóelőhívó hatását természetesen gátolhatják a beszédmegértés zavarai. Vizsgálatunk másik feladata annak feltárása, hogy az afáziás betegeknek a beszédfeldolgozás mely szakaszának deficitje akadályozza a fonémikus támpont felhasználását.

### **Módszerek**

Az afáziás betegeknek 30-50 feladatban vizsgáltuk részletesen mind a beszédmegértés, mind a beszédképzés folyamatait. Hangsúlyoznunk kell, hogy jelen tanulmányunkban post hoc elemezzük a fonémikus támpont jól kontrollált körülmények közötti hatását. Az egyes vizsgálatokat, amelyekkel ezt összevetjük, nem a jelen elemzés céljából végeztük. Tanulmányunkban e feladatok közül csak a kérdésfeltevésünk szempontjából relevánsnak kapott eredményeket ismer-tjük.

**A Boston Megnevezési Teszt (BMT).** A feladat 60 fekete-fehér vonalas rajzot tartalmaz, amelyeket a vizsgált személynek egyenként meg kell neveznie. Ha láthatóan felismerési problémával küszködik, egy definícióval segítjük (például a 'polip' képénél „tengeri állat”). Ha a vizsgált személy nem képes megnevezni a képet, akkor a szó

első hangjainak megkezdésével (fonémikus támpont) próbáljuk segíteni a megnevezést. Az egészséges személyekkel végzett vizsgálatokról kissé részletesebben korábban már beszámoltunk (Osmanné Sági 1998). A fonémikus támpont sikerességét (FTS) személyenként százalékként jellemeztük: a támponttal sikeresen előhívott szavak száma osztva az összes támponttal való próbálkozás számával és beszorozva százszal. Ezt az értéket hasonlítottuk össze az egyéb vizsgálatok eredményével. Ezek: Western Aphasia Battery (Kertész 1979; Osmanné Sági 1991) indexei (fluencia, megértés, ismétlés, megnevezés), illetve afázia kvóciense. Fonémapár (190 pár egy diszkriminatív jegyben – zöngésség vagy a képzés helye szerint – különböző értelmetlen szótag) diszkriminációs feladat; 120, egy hangban eltérő szópár-megkülönböztető feladat; szövmegértés, mondatmegértés (vö. Osmanné Sági 1994); szóismétlés (Osmanné Sági 1998).

Az intelligenciát a betegeknél a nem verbális Raven-teszttel mértük, amelyben geometriai jellegű ábrák közötti formai, illetve logikai kapcsolatokat kell megtalálni. E feladat teljesítése egyben feltételezi a vizuális észlelési folyamatok épségét.

Tekintettel arra, hogy nem minden egyes feladatot tudtunk minden egyes beteggel elvégezni, mellőznünk kellett a szofisztikáltabb statisztikai eljárások alkalmazását. Az egyszerűbb statisztikai korrelációs számításokat (illetve *t*-próbákat) minden egyes feladatnál külön-külön végeztük az érvényes betegszámmal.

### **Vizsgálati személyek**

A Boston Megnevezési Tesztet a képek felismerhetőségét vizsgáló előkísérletek, majd módosítások után összesen 173 egészséges felnőtt emberrel próbáltuk ki. A betegekkel végzett vizsgálatunkban 17 afázias vett részt. Az afáziát a betegeknél az agyi érrendszer megbetegedése okozta, általában 2-3 évvel vizsgálataink előtt. A betegek jobbkezese, kivéve T.T.-t. Többnyire csak a bal féltekéjüket érte a primer károsodás (a balkezes T.T.-nél a jobbót). A betegek többsége logopédiai terápiában részesült.

1. táblázat: Az afáziás betegek alapadatai

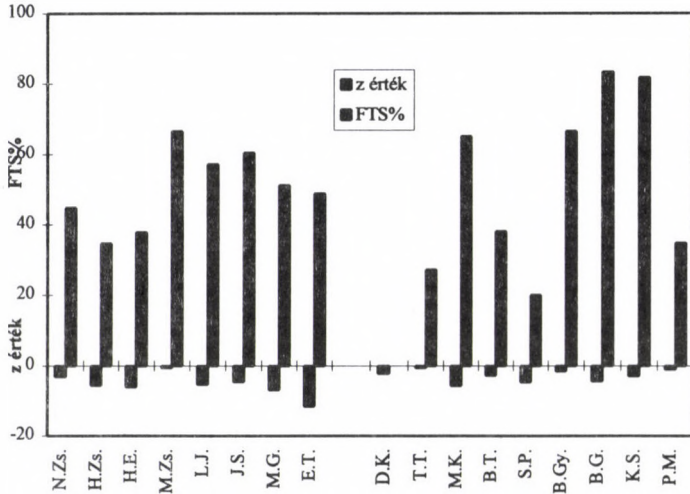
A személyre vonatkozó adatok					Afázia-típus	
név	kor	nem	iskola	etiológia		idő
N.Zs.	22	nő	ált.	stroke	Tr.motor.	2 év
H.Zs.	40	nő	ált.	stroke	Broca	2 év
H.E.	40	nő	közép	aneurysma	Broca	2 év
M.Zs.	43	nő	ált.	stroke	Broca	6 év
L.J.	44	nő	felső	aneurysma	Nonfluens	6.5 év
J.S.	47	férfi	ált.	stroke	Broca	2.5 év
M.G.	48	férfi	közép	aneurysma	Broca	1.5 év
E.T.	54	férfi	felső	stroke	Broca	3.5 év
D.K.	40	nő	közép	stroke	Wernicke	3 év
T.T.	58	férfi	közép	stroke	Wernicke	2 év
M.K.	64	férfi	felső	stroke	Wernicke	9 hó
B.T.	59	férfi	közép	aneurysma	Wernicke	1.5 év
S.P.	50	férfi	közép	stroke	vezetési	1 év
B.Gy.	57	férfi	ált.	stroke	vezetési	7 hó
B.G.	40	nő	felső	stroke	anomikus	10 hó
K.S.	52	nő	felső	aneurysma	anomikus	16 év
P.M.	58	nő	ált.	Stroke	anomikus	2 év

### Eredmények

A kapott eredmények elemzésekor első kérdésünk az volt, hogy a BMT-ben elért teljesítmény nem határozza-e meg önmagában a fonémikus támpont sikerességét (FTS). Ilyen összefüggést nem sikerült kimutatnunk. Az FTS nem korrelált szignifikánsan sem a nyerspontokban megadott teljesítménnyel, sem a saját egészséges kontrollcsoportjától z-értékben megadott eltéréshez viszonyítva (vö. 1. ábra). Ugyanakkor, míg a betegeknek a szavankénti teljesítménye korrelált



az egészségesekével, az FTS-ük nem mutatott a szavankénti teljesítményükkel korrelációt.

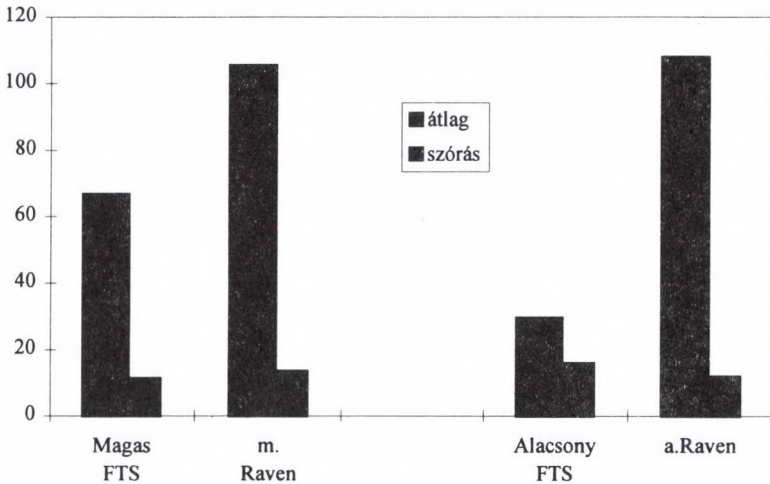


1. ábra

A BMT-teljesítmény és a fonémikus támpont kapcsolata

Mivel a BMT feltételezi a vizuális folyamatok épségét, szükségesnek véltük ez utóbbiakat is legalább szűrővizsgálat szintjén kontrollálni. Nyilvánvaló, hogy ha a betegek jó eredményt érnek el e feladatban, nem valószínű a vizuális folyamatok károsodása, míg alacsony értékeknél csak e teszt segítségével nem elkülöníthetők a vizuális észlelés és a nem verbális intelligencia deficitjei. A két érték az egész betegcsoport eredményeit véve összehasonlítási alapul, nem korrelál. Ha a Raven-teszt eredményei alapján két csoportra bontjuk a betegeket, külön-külön a magas, illetve az alacsonyabb Raven-értékek sem korrelálnak az FTS-sel. Akár a WAB fluencia alapján, akár a BMT nyerspont alapján, mediána mentén történő két csoportra bontásban vizsgáltuk a csoportok Raven-teszt eredményeit, ez utóbbiak nem mu-

tattak szignifikáns eltérést. Ha azonban az FTS értékek szerint bontjuk magas és alacsony értékű csoportra a betegeket (a két csoport FTS érték szerint egymástól szignifikánsan különbözik), a Raven-érték nagyságában nem különböznek szignifikánsan. Azonban az alacsony FTS csoportnál szignifikáns korreláció mutatható ki a FTS és a Raven IQ között.

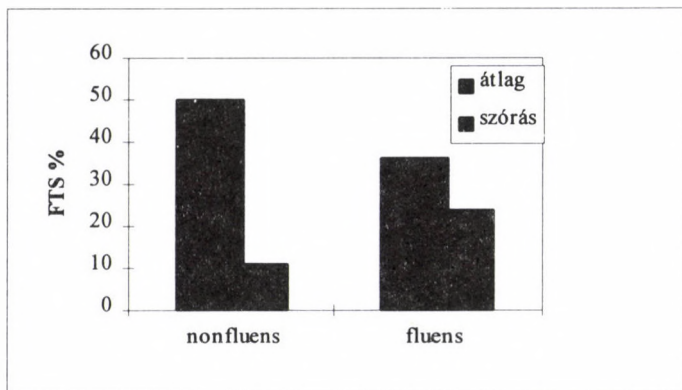


2. ábra

A fonémikus támpont kapcsolata a nem verbális Raven-teszt eredményeivel

A fonémikus támpont sikerességét összevetettük a Western Afázia Tesztnél kapott jellemzőkkel. Az afázia kvócienssel, amely az afázia súlyosságát jellemzi, nem kaptunk korrelációt. Ugyancsak nincs kapcsolat az FTS és a beszéd folyékonyági indexének nagysága, valamint a megnevezési tesztek eredményei között. A nem fluens és a fluens betegcsoportoknál a fonémikus támpont százalékban kifejezett sikerességét két mintás *t*-próbával is összehasonlítottuk. E módszerrel sem kaptunk szignifikáns különbséget a két csoport között, bár a nem

fluens csoportnál az FTS átlaga magasabb, mint a nonfluens csoportnál, viszont alacsonyabb az értékek szórása (vö. 3. ábra).



3. ábra

A fonémikus támpont sikeressége a két afáziás csoportnál

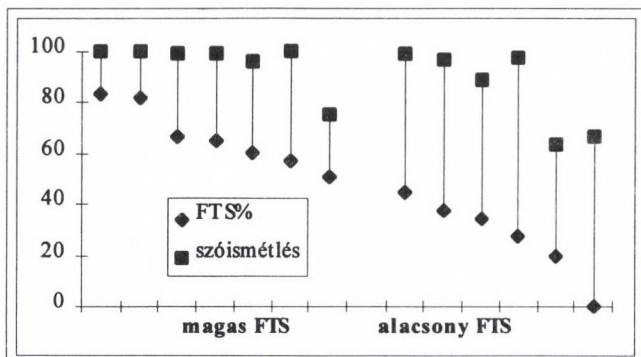
Korrelációt találtunk az FTS és a WAB megértési indexe ( $p < 0.05$ ), valamint az FTS és az ismétlési index ( $p < 0.01$ ) között. A továbbiakban az e folyamatokat részletesebben vizsgáló próbák eredményeit hasonlítottuk össze.

A fonémapár-diszkriminációval statisztikailag szignifikáns korrelációt nem találtunk, bár a kapcsolat tendencia szinten ( $p < 0.1$ ) jelentkezik. Azonban az FTS alapján magas és alacsony csoportok közül az alacsonynál szignifikáns korreláció mutatható ki ( $p < 0.05$ ). Ha a fonéma-diszkrimináció alapján bontjuk két csoportra az eredményeket, egyik csoport eredménye sem korrelál az FTS-sel.

A szópár-diszkriminációs feladat eredményei (4. ábra) és az FTS közötti korreláció viszont már statisztikailag is szignifikáns ( $p < 0.01$ ). Ha azonban szétválasztjuk az alacsony, illetve a magas teljesítményt mutató betegeket, akár a szópár-diszkriminációs teszt, akár az FTS alapján, a magas eredményt adóknál a két feladat eredményei nem



korrelálnak. Korreláció csak az alacsony eredményt elérteknél mutatható ki ( $p < 0.01$ ).



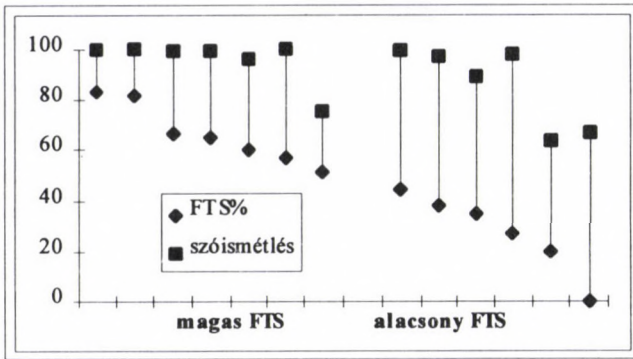
4. ábra

Az FTS eredményei és a szópár-diszkrimináció kapcsolata

A szómegértési feladat eredményei, amelyben 4, főnevet ábrázoló kép közül kell kiválasztani a hallott szónak megfelelő rajzot, szintén nem korreláltak az FTS-sel.

A mondatmegértést vizsgáló feladatok eredményei sem korreláltak a fonémikus támpont felhasználásának sikerességével (5. ábra), sem az elsősorban munkamemóriát vizsgáló Token-tesztben, sem pedig a beágyazott alárendelt mondatok megértését vizsgáló mondat/kép kiválasztásos feladatban.

Az ismétlési feladatban értelmes szavak és értelmetlen logatomok ismétlését vizsgáltuk. Mind a két ingertípusnál szignifikáns korrelációt kaptunk a fonémikus támpont sikerességével. Azonban, ha a betegeket az FTS nagysága szerint két csoportra osztottuk, és így hasonlítottuk össze az ismétlési teljesítménnyel, akkor csak az alacsony FTS értéket mutató betegeknél sikerült kimutatni mindkét ingertípusnál korrelációt. Ha viszont az ismétlési eredmények sikerességét vettük össze az FTS értékével, ilyen összefüggés nem volt kimutatható.



5. ábra

A fonémikus támpont sikeressége és az értelmes szavak ismétlése

Vizsgálatainkban a fonémikus támpont hatékonyságát illetően nem kaptunk statisztikailag szignifikáns különbséget a nem fluens és fluens betegek között. A 2. táblázatból jól látható, hogy mind a magas, mind az alacsony FTS értékű betegcsoportok felét a nem fluens (Broca, transzkortikális motoros) betegek alkotják. Az FTS a leghatékonyabbnak két anomikus betegnél bizonyult

A Wernicke- és a vezetékes afáziások kétharmadánál az FTS alacsony hatékonyságú volt, míg 2 betegnél magas. Tehát azt a klinikai benyomást, hogy a nem fluens afáziásoknál a fonémikus támpont hatékonyabb, mint a fluenseknél, vizsgálatunk ebben a formában nem támasztotta alá. Azonban a fluens Wernicke- és a vezetékes afáziás betegek többségénél valóban kevésbé hatékony a fonémikus szókezdés a szó előhívásában. Meg kell jegyeznünk, hogy az iskolázottság is eltért a két csoportnál, hiszen az egyetemet végzettek, egy kivétellel, a magas FTS csoportba kerültek. (Tisztázásra szorul, hogy az egészséges személyeknél kimutatható-e ilyen összefüggés.) Azonban e kérdésben az iskolázottság döntő hatása ellen szól, hogy az FTS alapján szétválasztott két csoport nem különbözik egymástól a Ravennel mért IQ alapján.

2. táblázat: Az FTS szerint csoportosított betegek afázia-típusai és iskolázottsága

Név	Iskola	Típus	FTS	Név	Iskola	Típus	FTS
B. G.	felső	anomikus	83.3 3	E. T.	felső	Broca	48.83
K. S.	felső	anomikus	81.8 1	N. Zs.	ált.	Tr. motoros	44.82
M. Zs.	ált.	Broca	66.6 6	B. T.	közép	Wernicke	38.09
B. Gy.	ált.	vezetési	66.6 6	H. E.	közép	Broca	37.83
M. K.	felső	Wernicke	65.2 1	P. M.	ált.	anomikus	34.78
J. S.	ált.	Broca	60.4 6	H. Zs.	ált.	Broca	34.69
L. J.	felső	nem fluens	57.1 4	T. T.	Közép	Wernicke	27.27
M. G.	közép	Broca	51.2 1	S. P.	Közép	vezetési	20
-	-	-	-	D. K.	Közép	Wernicke	0

A fonémikus támpont hatékonysága nem függ attól sem, hogy mennyire ép a tárgyak vizuális ábrázolásának a verbális lexikont aktiváló hatása. Ez utóbbi a szógyakoriságtól függ (ezzel korrelál az egészségesek szavankénti teljesítménye, ez utóbbival pedig a betegé), azonban az FTS nem mutat ilyen kapcsolatot.

Mennyiben függ az FTS a vizuális folyamatoktól? Speciális feladatok nem szerepeltek vizsgálati tesztanyagunkban, csak a nem verbális Raven-teszt és a Rey-Osterrieth Complex Figure Test (értelmetlen figura másolása, ill. emlékezetből való rekonstruálása.) Ez utóbbi feladattal az FTS semminemű korrelációt nem mutatott. Ahogy már említettük, a nem verbális intelligenciát mérő vizuális Raven-tesztben sem találtunk különbséget a két csoport eredményei között. Viszont az alacsony FTS értékű csoportnál az FTS szignifikánsan korrelált a Raven pontértékkel ( $p < 0.05$ ). A Raven-teszt megoldásakor verbális



stratégiákat is lehet használni, amely legalább a lemma szint aktiválásával járna. Ennek sikertelensége azonban megmutatkozna az IQ nagyságában is, tehát az implicit verbális folyamatok deficitje nem okozhatja a fenti korrelációt. Úgy véljük, azonos IQ-érték mellett ilyen korreláció talán a vizuális folyamatok valamely enyhe deficitjéből adódhat. (Ennek bizonyításához speciális vizsgálatok szükségesek.)

Vizsgálatunkban központi helyet foglal el annak a kérdésnek a tisztázása, hogy a beszédmegértés zavarai mennyiben okozhatják a fonémikus szókezdés alacsony hatékonyságát. Az FTS és a WAB megértési indexe közötti szignifikáns korreláció egyértelműen igazolja egy ilyen hatás fennállását. (Jelen közleményünkben nem elemezzük a probléma és az afázia típusa közötti kapcsolatokat.) További kérdés, hogy a megértés mely szakasza érintett. E kérdés megválaszolása céljából vizsgáltuk az egy diszkriminációs jegyben különböző szótagpárok diszkriminációját. Csak az alacsony FTS csoportnál volt kimutatható szignifikáns kapcsolat az FTS és a fonéma-diszkrimináció eredményei között. Ugyanakkor, ha a fonéma-diszkriminációs teljesítmény alapján osztottuk két csoportra a betegeket, egyik csoport eredményei sem korreláltak. Tehát a fonémikus támpont hatékonyságát nem határozza meg önmagában a fonéma-pár diszkrimináció, bár az alacsony hatékonyságúaknál a diszkrimináció is alacsony.

Ha az egy jegyben különböző szópárok diszkriminációjánál a teljes betegcsoport eredményét vetettük össze a FTS-sel, a korreláció szignifikánsnak bizonyult ( $p < 0.01$ ). Ha akár a szódiszkrimináció, akár az FTS alapján szétválasztjuk a mediána mentén a magasán és alacsonyan teljesítő csoportokra a betegeket, az alacsonyan teljesítőknél lehet mindkét feladatra bontás szerint szignifikáns korrelációt találnunk ( $p < 0.01$ ). A szó/kép összehasonlítási feladat teljesítményei nem korrelálnak az FTS-sel, ahogy a mondatmegértési feladatok eredményei sem. Tehát a megértési feladatok közül az alacsony fonéma-diszkrimináció, szópár-diszkrimináció és a fonémikus támpont hatékonysága között állapítható meg egyértelmű kapcsolat.

A szóismétlési feladat eredményei is korrelálnak az FTS-sel, de csak az alacsony hatékonyságot mutató csoportnál, mind az értelmes

szavaknál, mind a logatomoknál. Az ismétlés eredményei alapján történő csoportra bontás eredményei viszont nem mutatnak kapcsolatot az FTS-sel. Ha viszont az FTS alapján szétválasztott csoportoknál vizsgáljuk a szópár-diszkrimináció és a szóismétlés kapcsolatát, a magas eredményt elért csoportnál nem mutatható ki kapcsolat, míg az alacsony csoportnál a két feladat eredményei értelmes szó és logatom ismétléskor egyaránt szignifikánsan korrelálnak a diszkrimináció eredményeivel. Külön vizsgáltuk a szópár-diszkrimináció és a szóismétlés kapcsolatát az ismétlés alapján két csoportra bontva a beteg populációt. Egyértelmű korrelációt a két feladat eredményei között csak az alacsony ismétlési eredményeket adó betegeknel mutattunk ki mind az értelmes szavaknál, mind a logatomoknál. Ha viszont a szópár-diszkrimináció alapján bontottuk magas és alacsony teljesítményű csoportokra a betegeket, szignifikáns korrelációt a hozzájuk kapcsolódó ismétlési értékekkel nem mutattunk ki. Azaz, az alacsony ismétlési teljesítményt elért betegek a diszkriminációs tesztben is alacsony teljesítményt érnek el, de az alacsony diszkriminációs teljesítmény nem feltétlenül jár alacsony ismétlési teljesítménnyel.

Összegezve: vizsgálataink nem igazolták, hogy a nem fluens betegeknel a fonémikus támpont hatékonyabb lenne a szóelőhívásban, mint a fluens afáziásoknál. Azonban a fluens afáziások közül a Wernicke- és a vezetésés afáziások kétharmadánál a fonémikus támpont hatékonysága alacsony, 50% alatti. A nem verbális intelligencia nem befolyásolja az FTS-t, azonban az alacsony FTS értékű csoportnál korrelálnak. E jelenséget talán a vizuális feldolgozás, illetve a vizuális és verbális folyamatok közti kapcsolat zavarának lehet tulajdonítani. A beszédmegértési feladatok közül a fonémikus támpont hatékonysága mind a fonémapár-, mind a szópár-diszkriminációs feladatban jelentkezik, amelyek károsodása érintheti az ismétlési feladat eredményeit is. Az egyszerű tárgyakat ábrázoló képek megnevezésekor a kép maga részlegesen vagy teljesen aktiválhatja a szó mentális lexikonjának lemma szintjét. Ha e szint nem képes a lexéma szint teljes mértékű aktivációjára, a szókezdet ebben segítséget nyújthat. A fonémikus támpont hatékonyságát befolyásolhatja a beszédhang feldolgozás,



pontosabban a fonéma feldolgozás zavara, jelentősebben akkor, ha szavak kontextusában hangzanak el. Azaz a fonéma szint és a bemeneti lexikon zavarainak interakciója gátolja meg, hogy az önmagában értelmetlen, lexikális egységnek nem tekinthető szókezdet aktiválja a kimeneti lexikont (és az általa vezérelt posztlexikális tervező és kivitelező szinteket). Ehhez hozzájárulhat a vizuális, illetve vizuális/verbális folyamatok zavara, amely a lemma szint korlátozott aktivációját okozhatja.

### Irodalom

Caplan, D.: *Language. Structure, Processing and Disorders*. MIT Press. Cambridge, Mass. 1992.

Caramazza, A.: How many levels of processing are there in lexical access? *Cognitive Neuropsychology* 14. 1997, 177-208.

Goodglass, H.–Kaplan, E.–Weintraub, S.–Ackermann, N.: The „tip of the tongue” phenomenon in aphasia. *Cortex* 12. 1976, 145-153.

Gósy Mária: A szó felismerése: folyamatok és stratégiák. In: *Beszédkutatás '97. Szófonetikai vizsgálatok*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1998, 63-117.

Kertesz, A.: *Aphasia and Associated Disorders*. Grune and Sraton. New York 1979.

Levelt, W. J. M.–Roelofs, A.–Meyer, A. S.: A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22. 1999, 1-75.

Marslen-Wilson, W. D.: The cohort model. In: Marslen-Wilson, W. D. (ed.): *Lexical Representations and Process*. Cambridge Press. New York 1989, 75-108.

Osmanné Sági Judit: Az afáziák klasszifikációja és diagnosztikája I-II. *Ideggyógyászati Szemle* 44. 1991, 339-362.

Osmanné Sági Judit: A beszédmegértés vizsgálata afáziás betegeknél. In: *Beszédkutatás '94. Tanulmányok az elméleti és alkalmazott fonetika köréből*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézete. Budapest 1994, 121-139.

Osmanné Sági Judit: Szóismétlési zavar elemzése afáziás betegeknél. In: *Beszédkutatás '97. Szófonetikai vizsgálatok*. Szerk.: Gósy Mária MTA Nyelvtudományi Intézete. Budapest 1998, 142-169.

Készült a T 18391 sz. OTKA-támogatás keretében.





Készült: Regiszter Kiadó és Nyomda Kft.





